THE RELATIONSHIP BETWEEN TRANSGENE COPY NUMBER AND VARIEGATED EXPRESSION IN MICE

Gráinne Barkess

Ph.D.
University of Edinburgh
2001

Science has 'explained' nothing;

the more we know

the more fantastic the world becomes

Aldous Huxley

ABSTRACT

Pronuclear microinjection to produce transgenic animals is now a widespread technique. However, with this method there is no control over where in the genome the transgene integrates or the structure of the array produced. It has long been known that the position within the genome can have a profound effect on transgene expression. Evidence that the copy number also has an effect has recently emerged. To test the hypothesis that high copy number arrays can be responsible for variegation, site-specific recombination was employed to reduce the copy number at the same genomic location, allowing a direct comparison of different copy number arrays at the same integration site.

A BLG-loxP transgene was microinjected to produce transgenic mice. High copy founders were used to establish lines and their expression profiles were analysed using *in situ* hybridisation, Northern blots, and milk protein composition. Two out of five lines showed variegated expression with discreet patches of cells expressing within the mammary gland.

The variegating lines and one uniform line were then bred to a line of mice expressing a BLG-Cre recombinase transgene in the mammary gland. The double transgenic animals were analysed for a mammary-specific reduction in copy number and their expression patterns were also studied. In all three lines, after the reduction of copy number there had been a reduction in the number of cells in the mammary gland that expressed the BLG transgene. This was contradictory to the hypothesis that suggested that the variegation should be relieved by a reduction in copy number.

The same BLG-loxP lines were then microinjected with a PGK-Cre recombinase construct to produce a reduction of copy number early in development. Microinjected animals were analysed for reduced copy arrays, of which only one line showed evidence. These animals were bred to establish lines with reduced copy number arrays throughout the animal. Their expression profile was analysed as before. Animals that had a reduction to one copy showed no expression, while animals that showed a reduction to two copies showed limited patches of expression. This result mirrored that of the mammary-specific reduction.

BLG transgenes may require a buffer zone provided by high copy arrays to allow some transgenes within the array to escape genome effects. When the copy number is reduced, the percentage of cells that can escape the silencing is also reduced. It is therefore clear that in some cases, transgenes may only efficiently express as a multicopy array.

ACKNOWLEDGEMENTS

I would like to thank my supervisor Prof. John Clark for his support and encouragement throughout my PhD, and my University supervisor, Dr.Andrew Smith for productive discussions and technical help. I would also like to thank Dr. Bruce Whitelaw for helpful discussions and ideas.

I would like to thank the people in the lab: Claire Neil for putting up with me and answering my many questions, Wei Cui, Doug Strathdee, Roberta James, John Webster and Romi Pena for discussions and technical tips. Thanks to the staff at the SAU for looking after all the mice, Caroline McCorquodale for statistical advice, and the Graphics team for help with photography. Thanks also to Ray Ansell, Janie and Viv for cleaning up all that mess, and Tom Skinner and Jo Rowe for their encouraging remarks (!) while I was writing up. Thanks to all the PhD students: Kathirvel, Simon, Ricardo, Margaret, Stephen, Gillian, Yasmin and AnneMarie for livening up the office.

I would like to thank my parents and sister for encouraging me throughout my PhD. If they hadn't been as supportive I would have never finished. This can be the second book for Gallus Publishing! Thanks also to my friends: Barry, Kerrie-Ann, Sheila, Gareth and Pamela who knew how to distract me with the occasional night out.

I was financially supported by a BBSRC studentship and the Roslin Institute.

To my family and friends

CONTENTS

CHAPTER TWO: MATERIALS AND METHODS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS AND CONVENTIONS

The following abbreviations are used throughout this thesis:

AATB	alpha anti trypsin
APS	ammonium persulphate
BLG	β-lactoglobulin
bp	base pair(s)
BSA	bovine serum albumin
CAT	chloramphemicol acetyltransferase
cDNA	complementary DNA
CpG	cytidine guanosine dinucleotide
cm	centimetre
c.v.	coefficient of variance
CPM	counts per minute
dATP	2'-deoxyadenosine 5'-triphosphate
dCTP	2'deoxycytidine 5'-triphosphate
DEPC	diethyl pyrocarbonate
dGTP	2'-deoxyguanosine 5'-triphosphate
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	2'-deoxy (N) 5'-triphosphate
DPM	disintegrations per minute
DTT	dithiothreitol
dTTP	2'-deoxythymidine 5'-triphosphate
EDTA	ethylenediamine-tetra-acetic acid

ES cells	embryonic stem cells
EtBr	ethidium bromide
g	gram(s)
G1	first generation
HDAC	histone deacetylase
HP1	heterochromatin protein 1
HPRT	hypoxanthine-guanosine phosphribosyl transferase
HSV TK	herpes simplex virus thymidine kinase
kb	kilobase (1000bp)
kDA	kiloDalton
l	litre(s)
LacZ	meta-galactosidase gene
LB	Luria Bertani medium
LCR	locus control region
M	Molar (moles/litre)
µg	microgram(s)
µl	microlitre
µm	micron (one thousand of a millimetre)
mg	milligram(s)
min	minute(s)
ml	millilitre(s)
mM	millimolar
MIP	methylation induced premeiotically
MOPS	3-N-(morpholino) propanesulfonic acid
mRNA	messenger RNA
MUP	major urinary protein
NLS	nuclear localisation signal
ng	nanogram(s)

nm	nanometre(s)
NP40	octylphenoxypolyethoxyethanol
nt	nucleotide
OD	optical density
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PEV	position effect variegation
PFA	paraformaldehyde
PGK	3-phosphoglycerate kinase
PMSF	phenylmethylsulfonyl fluoride
rATP	adenosine 5'-triphosphate
rCTP	cytidine 5'-triphosphate
rGTP	guanosine 5'-triphosphate
RIGS	repeat induced gene silencing
RIP	repeat induced point mutation
RNA	ribonucleic acid
RNase	ribonuclease
rpm	revolutions per minute
rUTP	uridine 5'-triphosphate
s.d.	standard deviation
SDS	sodium dodecyl sulphate
sec	second(s)
SSC	standard saline citrate
	(0.15M NaCl,15mM tri-sodium citrate, pH7.0)
TAE	Tris-acetate electrophoresis buffer
TBE	Tris-borate electrophoresis buffer
TCA	trichloroacetic acid

TE	10mM Tris-HCl pH7.9, 1mM EDTA
TEA	triethanolamine
TEMED	N, N, N', N'-tetramethylenediamine
temp	temperature
TESPA	3-aminopropyltriethoxysilane
Tris	2-amino-2-(hydroxymethyl)-1,3-propandiol
tRNA	transfer RNA
Tween 20	polyoxyethylenesorbitan monolaurate
UV	ultraviolet
V	volt(s)
v/v	volume / volume
w/v	weight / volume
YAC	yeast artificial chromosome
°C	degrees Celsius

CHAPTER ONE

INTRODUCTION

Recent advances of recombinant DNA technology and embryo culture techniques have underpinned an explosion in the production of genetically manipulated higher eukaryotes. The introduction of foreign genes, termed 'transgenes', into the germline of organisms has enabled researchers to investigate the expression of foreign genes and their products. These transgenic organisms exhibit unique phenotypes and these characteristics are inherited by their offspring. The ability to alter the genome of organisms has opened up new fields of study and transgenic organisms provide unique opportunities to address many fundamental biological questions.

The use of transgenics has become widespread; thousands of transgenic mice, plants and livestock now exist (Old and Primrose 1994). Transgenics have been established to provide improvements in commercial characteristics of livestock (Bullock *et al* 1997). Improvements to crops have allowed alterations to farming methods, improving efficiency and reducing losses (Gressel 2000). Transgenic livestock have been produced that express biomedically important proteins in their milk, which can be used in treating human disease (Rudolph 1999). Transgenic mice have been used as models of human disease to help gain understanding into the causes of the diseases (Petters and Sommer 2000). They have also been used to study biological questions of gene function and developmental pathways.

In some cases transgenic mice showed an appropriate expression profile for the transgene used. Transgenes such as herpes thymidine kinase were found to express in the liver (Brinster *et al* 1981) and β-globin transgenes were found to express in erythroid tissues (Chada *et al* 1985). However, it has become apparent that there are numerous examples of unstable or poor transgene expression. Both the genomic integration site and the structure/copy number of transgenic arrays can lead to unstable expression (Dobie *et al* 1997). Work has been carried out that implicates high copy number in the silencing of transgenes, leading to variegated expression (Dorer 1997). This has profound implications for those researchers wishing to use this technology. A clearer understanding, therefore, of transgene regulation is required to allow this technique to be fully exploited.

This chapter reviews the historical development of techniques used to produce transgenic animals and then focuses on how these transgenes express. Initially, the mechanisms by which transgenes integrate into the genome and their array structure are discussed. The roles of factors known to influence expression (i.e. intrinsic factors such as enhancers and promoters, and extrinsic factors such as the position within the genome) are examined. Furthermore, aberrant examples of transgene expression and transgene silencing are also presented. The mechanisms by which transgenes are silenced show similarities to the broader field of epigenetics, and common factors are considered. These issues are all reviewed in relation to transgene expression in the mammary gland, which is the tissue studied in this project. The chapter concludes with the aims and approaches of the project, which is to investigate the relationship between the copy number of transgenes and variegated expression.

1.1 DEVELOPMENT OF TRANSGENIC TECHNIQUES

1.1.1 First production of transgenic animals

The ability to introduce genes into the germline of mammals is one of the most important recent advances in biology. The first transgenic mice were produced by microinjecting SV40 viral DNA into the blastocoel cavity of early embryos, which were then implanted into foster mothers. Some cells in the embryo incorporated the DNA, resulting in mosaic adult animals (Jaenisch and Mintz 1974). This work did not test if there was germline transmission of the integrated DNA.

The next attempts to produce transgenic animals involved infecting mouse embryos with retroviruses early in development. This resulted in the stable integration of Moloney leukemia retroviral DNA into the germ line (Jaenisch 1976). Several disadvantages exist with this method. The integration of the foreign DNA occurred at different stages of embryonic development resulting in mosaic animals. This method often generated animals that had more than one transgene integration site, because more than one cell in the early embryo was infected with the virus (Palmiter and Brinster 1985). Due to the occurrence of multiple integration sites, lines of mice had to be out-bred to segregate the various loci to produce single insertion lines.

1.1.2 Generation of transgenic animals by pronuclear injection

In 1980 an extremely important advance in transgenic technology was reported. Gordon and co-workers described a technique whereby naked DNA was injected into the pronucleus of a newly fertilised mouse oocyte (Gordon *et al* 1980). This method was to become the standard means to manipulate the mouse genome for sometime. The experiment showed it was possible to transfer foreign genes directly into mouse embryos. The foreign DNA integrated into the genome and was inherited by offspring of the founder transgenic animal. Injection of one-cell embryos was important for obtaining early integration, so allowing the foreign DNA to contribute to the genome of all somatic cells and importantly, the germline. The discovery of pronuclear injection as a new means of altering the genome of animals spawned a revolution in the analysis of gene expression (Houdebine 1997).

The efficiency of pronuclear injection to generate transgenic animals is determined by a number of factors. Early studies by Brinster and co-workers (Brinster *et al* 1985) investigated the effects of a number of parameters on the efficiency of integration, but not on transgene expression. These early experiments established that the most critical factors affecting DNA integration were the concentration and form of the DNA injected. Integration efficiency improved as DNA concentration increased, with an optimal concentration of one ng/μl; large amounts of DNA however, were toxic. Comparison of cytoplasmic versus nuclear injection showed the efficiency of DNA integration after cytoplasmic injection was very low. No significant difference in integration efficiency was found when the male or female pronucleus was injected. However, in mice the male pronucleus is

favoured, being the larger of the two nuclei in this species. DNA molecules were injected in four different forms; supercoiled, restricted and blunt ended, restricted to form 'sticky ends', or restricted to form dissimilar ends. Linear DNA integrated five fold more efficiently than supercoiled DNA; it was also reported by these workers that linear DNA molecules with similar or dissimilar ends may integrate more efficiently than linear molecules with blunt ends. Hybrid strains of mice (i.e. C57 x SJL) were eight fold more successful than inbreds (i.e. C57) (Brinster *et al* 1985). This is likely to be due to the overall improvement in vigour of hybrid animals compared to inbred animals.

Pronuclear injection is currently the most commonly used technique for the generation of transgenic animals. It is however, an inefficient process. Of the embryos that survive the injection procedure only ~10% produce offspring, of which ~40% are transgenic (Boyd and Samid 1993). The success rate is variable between laboratories and between individual researchers due to the highly skilled nature of the technique.

1.1.3 Production of transgenic livestock by pronuclear injection

The technique of pronuclear injection has been used to generate transgenic livestock including pigs, cattle and sheep. It is less efficient in these species when compared to transgenic mice (Pursel and Rexford 1993, Maga and Murrey 1995, Eyestone 1994). In addition to the problems associated with the integration of new genes, there are inefficiencies associated with collecting and culturing fertilised eggs, as well as embryo transfer in livestock animals. Due to the increased generation time, smaller

litter size and the extra cost of caring for livestock, transgenic studies in livestock species are a major undertaking financially and logistically (Wall and Seidel 1992). Early studies investigated the use of transgenes controlling the productivity of the animal i.e. growth hormone transgenes to increase the rate of growth and the ratio of feed: growth (Pursel and Rexroad 1993). These early studies showed the problems of pronuclear injection in relation to the control of gene expression. Levels varied between lines and transgene expression could be low. Because of the poor efficiency of pronuclear injection and lack of control of gene expression, other methods for transgenesis were sought. Later experiments developed biomedical applications such as bioreactors for the production of therapeutically important proteins (Houdebine 2000). Transgenic livestock are also being used to investigate methods to make xenotransplantation more viable (Logan 2000).

1.1.4 ES Cells

The ability to grow pluripotent cells in culture has been central to developing other methods of transgenesis. So-called embryonic stem (ES) cells were derived from the inner cell mass of the mouse blastocyst. The cells could be maintained in an undifferentiated state in culture and displayed a normal karyotype (Evans and Kaufman 1981). If the ES cells remain pluripotent, after their return to recipient blastocysts, the cells can contribute to the formation of all tissues, including the germline (Bradley *et al* 1984).

ES cells can be manipulated *in vitro* without affecting their ability to contribute to the germline. It was demonstrated that ES cells could still

generate germline chimeras after transfection with a resistance marker gene. Neomycin phosphotransferase genes from bacterial transposons confer resistance to aminoglycoside antibiotics, allowing a selection strategy to be used (Berg 1981). ES cells, therefore, provided an exciting new method for introducing specific genes into the mouse genome.

Exploiting homologous recombination in murine ES cells allows integration of transgenic sequences into the genome at a controlled site (Capecchi 1989). Point mutations, deletions and gene inactivation have all been achieved by gene targeting in ES cells (Jasin *et al* 1996). Once the modification of the genome has been carried out in the ES cell, it is then injected into mouse blastocysts where it can contribute to the embryo producing a chimeric animal. If the ES cell has also contributed to the germline, the chimeric animals are used to establish lines of mice with the targeted alterations in their genome. ES cell transgenesis has now been used to produce mice with controlled sites of integration and single copy arrays of the transgene of interest (Bronson *et al* 1996). ES cell manipulation in mice has allowed targeted integration of transgenes, allowing better control over the expression profile (Jasin *et al* 1996). At present, no ES cells are available for livestock species.

1.1.5 Nuclear transfer

Recent advances in nuclear transfer have expanded the means by which genetic modification of animals other than mice can be carried out. Nuclear transfer has now been achieved in a number of species including sheep, mouse, goat and cattle (Clark *et al* 2000). Nuclear transfer involves

the enucleation of an oocyte followed by the electrofusion of this cell with a donor cell. The donor cell can be somatic in origin. This allows for easier manipulation of livestock genomes, circumventing the need for ES cells in these species (Wilmut *et al* 1997). It is likely in the future that genetic modification to livestock animals will be carried out in somatic cells for subsequent use in nuclear transfer. Currently this route is not as well established as ES cell targeting. The first gene targeted sheep by nuclear transfer has been reported whereby the ovine α-1(I)procollagen locus was targeted with a human α1-antitrypsin (AAT) contained within a BLG expression vector, which was designed to direct expression of the AAT to the mammary gland (McCreath *et al* 2000). Other recent work produced lambs with a deletion in the PrP gene, which codes for prions, the causal agent in spongiform encephalopathies (Denning *et al* 2001).

1.1.6 Transgene manipulation

1.1.6.1 Cre recombinase and loxP sites

Cre recombinase (Cre stands for causes recombination) is a member of the integrase family of recombinases along with FLP from *Sacchromyces cerevisiae* (Kilby *et al* 1993). Cre recombinase from bacteriophage P1 is a 38kDA protein and it exists as a monomer in solution (Sauer 1993, Mack *et al* 1992). Cre recombinase mediates efficient intramolecular (excisive or inversional) and intermolecular (integrative) site-specific recombination between target recognition elements known as loxP sites. This system has now been used in a wide variety of experiments, and is a routine method for the manipulation of transgenes (Metzger and Feil 1999).

LoxP sites are 34bp long and consist of two inverted repeats surrounding an 8bp asymmetric spacer region. Cleavage and strand exchange of recombining loxP sites occurs within this spacer region. The loxP sites exhibit directionality as a consequence of the asymmetry of the spacer region (Hoess and Abremski 1984). If two loxP sites are integrated into the same DNA molecule in a direct orientation then excision occurs leaving one loxP site behind (figure 1.1). If the loxP sites are in opposite orientations then inversion of the intermediate DNA occurs.

A single molecule of Cre is bound to each inverted repeat of the loxP site (Mack *et al* 1992). Prior to strand exchange Cre recombinase cleaves the DNA to generate 6bp staggered cuts. The single stranded ends formed must then pair with the complementary single strand sequence of the other loxP partner in recombination (Hoess *et al* 1986). The crystal structure of Cre recombinase bound to a loxP site has been determined. The synapsed structure of four Cre recombinases and two loxP sites resembles the models of four way Holliday junction intermediates (Guo *et al* 1997).

Cre mediated recombination proceeds efficiently with supercoiled and linear DNA molecules (Sauer and Henderson 1989). In *Sacchromyces cerevisiae* the ability of Cre recombinase to perform precise recombination events on chromosomes is unimpaired by chromatin structure. However, Baubonis and Sauer (1993) found that in human osteosarcoma cells chromosomal position effects had a strong influence on the rate of targeting to a chromosomal loxP site. Cre recombinase is also very specific; the introduction of FLP recombinase will not recombine loxP sites, neither will Cre recombine FLP binding sites, which are similar to loxP sites (Sauer 1987).

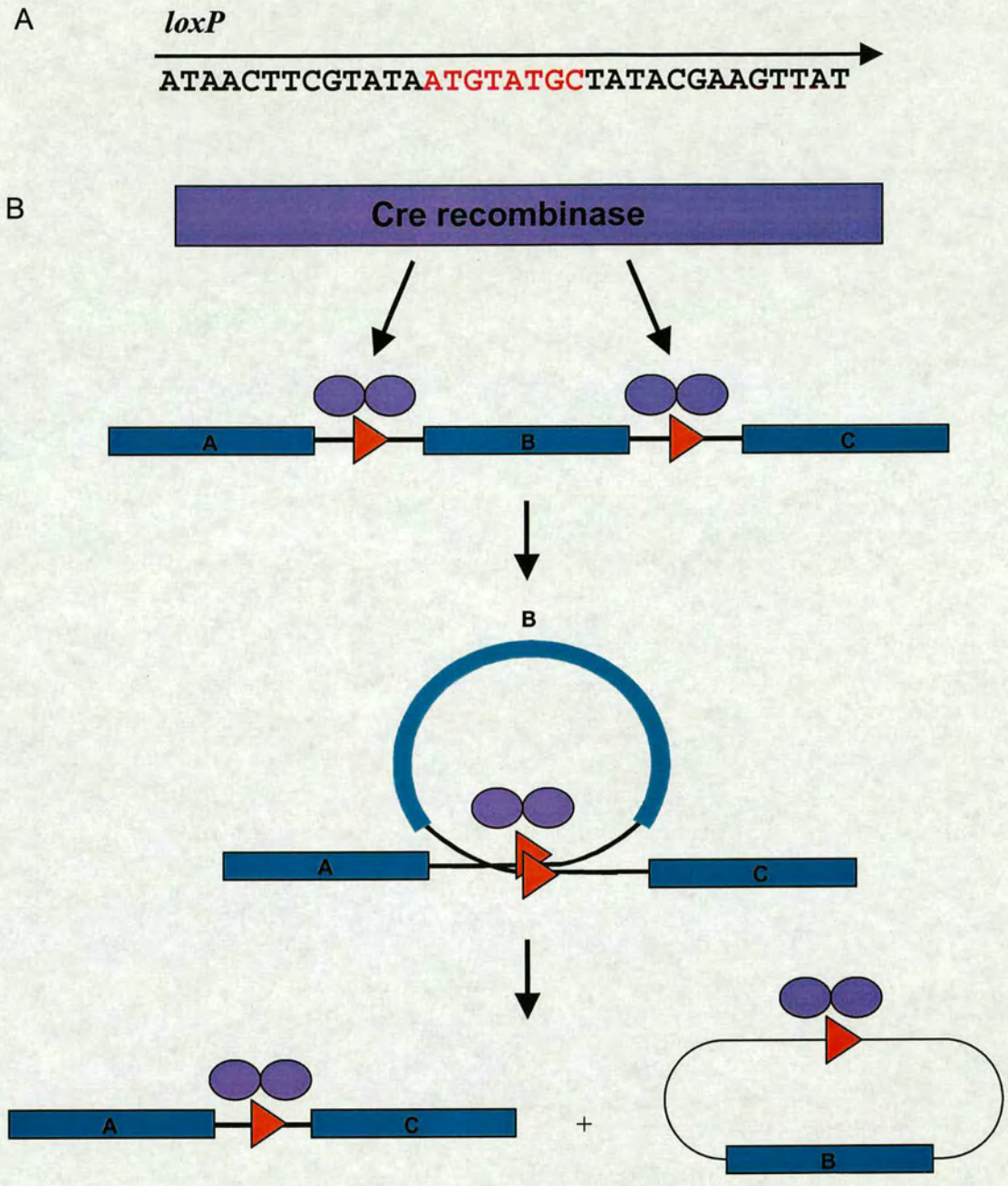


Figure 1.1. Cre-loxP recombination. **(A)** The sequence of the loxP site, the spacer region is red. **(B)** Cre catalysed excision. Two Cre molecules bind to each loxP (red triangle). Intervening DNA between adjacent loxP sites is removed, leaving one loxP site still integrated in the genome. The reaction can occur in reverse.

The length of the loxP target size was thought to make it unlikely to occur at random in higher eukaryotic genomes, so preventing the possibility of recombination with an endogenous eukaryotic sequence (Kilbey *et al* 1993). However, it has been shown that sequences from mammalian DNA that are divergent from the traditional loxP site can support Cre mediated recombination in mammalian cells (Thyagarajan *et al* 2000). Until recently no reports of loxP site infidelity in transgenic animals existed. However, Schmidt *et al* (2000) reported that expression of Cre driven by the protamine I promoter in postmeiotic spermatids led to Cre catalysed chromosomal rearrangements in the absence of loxP sites. All males derived from these transgenic mice were infertile due to abortion of early embryos whose chromosomes had been rearranged. The authors suggested that if chromosomal rearrangements were occurring in somatic cells that it may not be as easily detected. This report emphasised the removal of Cre after use and some self-excising vectors are now available (Bunting *et al* 1999).

1.1.6.2 Applications of Cre recombinase

The advantage of Cre-loxP is that host-specific functions are not required for their activity. Cre recombinase has been shown to catalyse site-specific recombination in the genome of yeast and eukaryotic cultured somatic and ES cells, and has been shown to catalyse site-specific recombination *in vivo* in transgenic animals (Sauer 1987, Sauer and Henderson 1989, Fukushigi and Sauer 1992, Ludwig *et al* 1996). It is now used as a powerful tool for genetic engineering in a wide variety of systems.

The ability to excise precise pieces of DNA can be used for different purposes. For example, transgenes can be activated by removing an

intervening stop sequence between the promoter and coding region of the transgene (Lasko *et al* 1992). It has been shown that Cre recombinase could be used in mice at high efficiency to excise an entire transgene *in vivo* or to reduce multicopy arrays to single copy. The authors suggest the frequency and extent of excision may depend on a variety of factors including the size of the fragment to be excised, the genomic site of Cre recombinase action, and the developmental point of Cre action (Orban *et al* 1992, Lakso *et al* 1996). Deletions of at least 200kb of DNA are possible when a Cre expressing plasmid is electroporated into ES cells (Li *et al* 1996). Other groups have reported the generation of deletions of up to 3 - 4 centimorgans using Cre-loxP (Ramirez-Solis, Liu and Bradley 1995). Cre expressing mice have also been used to produce cell type specific gene inactivation (Gu *et al* 1994). Alternatively Cre-loxP can be used to target a transgene to a specific genomic site containing a loxP site. This catalyses the reverse reaction, whereby a fragment of DNA carrying a loxP site can be targeted to the genomic loxP site by intermolecular recombination (Kühn and Schwenk 1997)

1.1.6.3 Expression of Cre recombinase in different tissues

Cre recombinase has been expressed specifically in the mammary gland of transgenic mice (Wagner *et al* 1997, Selbert *et al* 1998). Cre controlled by the mouse mammary tumour virus (MMTV) long terminal repeat (LTR) showed recombination in all tissues and was not specific to mammary alone. However, Cre controlled by the whey acidic protein (WAP) or beta-lactoglobulin gene was expressed in a tissue-specific manner. The temporal pattern of expression mirrored that of the promoters used to drive Cre recombinase.

A number of groups have generated Cre recombinase transgenic mice strains that mediate deletion of floxed genes (alleles surrounded by two loxP sites) very early in development i.e. preimplantation embryo or in oocytes. One report described the expression of Cre recombinase under the control of adenovirus EIIa promoter targeting expression to the early mouse embryo. This generated genetic changes to floxed alleles in all cells of the developing animal including the germ cells (Lakso *et al* 1996).

Phosphoglycerate kinase (PGK) is expressed in all somatic cells (Adra *et al* 1987). The PGK promoter has successfully been used to drive early and uniform Cre expression. When PGK Cre females are bred to males carrying a floxed reporter gene, recombination occurs even in offspring that do not inherit the PGK Cre gene. This suggested that the PGK Cre was already active in the diploid oogonium and that during meiosis the Cre transcript or protein was transmitted to both Cre^- and Cre^+ daughter cells (Lallemand *et al* 1998).

Transient expression of Cre has been achieved by pronuclear microinjection of either a circular Cre plasmid vector or cytoplasmic injection of an *in vitro* transcribed Cre mRNA (Araki *et al* 1995, de Wit *et al* 1998). Both methods led to a very high level of recombination and avoided the integration of the Cre construct. This is a rapid method for Cre mediated recombination and avoids having to breed two transgenic lines together.

1.2 STRUCTURE OF TRANSGENE LOCI

1.2.1 Analysis of transgene integration

With pronuclear injection neither the site of integration, nor the number of copies of foreign DNA can be controlled. The integration site is thought to occur randomly throughout the genome. Analysis of transgenic loci produced by pronuclear injection indicates that integration of a single transgene is a rare event: transgenes generally integrate as an array arranged in a head to tail manner (Wall and Burdon 1997). The number of copies within the array can range from one to several hundred (Brinster *et al* 1981, Palmiter *et al* 1982). This inability to control where in the genome and the number of copies that integrate is a drawback to the use of pronuclear injection.

1.2.2 Production and integration of concatemers

A multi-step model for the generation of transgene arrays has been proposed whereby an array of introduced genes is formed first, and then the array becomes incorporated into the genome at a break in the chromosome (Bishop and Smith 1989). It has been suggested that tandemly arranged concatemers of the input DNA are built up by a process of homologous, but non-conservative recombination. This predicts that a circularly permutated array is formed before integration (Bishop and Smith 1989). Recent studies have shown that some transgenes can integrate with nearly intact 5′ and 3′ copies (Cranston *et al* 2001). This data, therefore, supports the theory that

injected transgenes concatermize to form multiple arrays and that a linear concatemer is then integrated into the genome.

High levels of mosaicism are often observed in transgenic founder animals indicating that the DNA must have integrated after the fertilised oocyte has divided (Whitelaw *et al* 1993). It may be that integration occurs into a replication loop, which may explain some of the duplications, insertions of flanking or filler sequence found at transgene integration sites (Bishop 1997).

Integration appears not to be affected by chromosome morphology (i.e. acrocentric, metacentric or dicentric) or chromosome position even in the presence of homologous sequence in the genome (Dellaire and Chartrand 1998). Hamada *et al* (1993) proposed that the integration of the transgenic array into the host genome occurs by end ligation of the transgene concatemer and chromosomal breakage, possibly stabilised by limited two to three nucleotide homologies occurring between the transgene and genomic ends. Indeed, some studies have found only very small regions of homology of 1-6 nucleotides at the junctions between transgenic and genomic sequence (McFarlane and Wilson 1996).

1.3 EXPRESSION OF TRANSGENES

A transgenic experiment utilising pronuclear injection usually involves the generation of several lines containing the same transgene. Each transgenic founder is the result of a unique set of molecular events, and the integration site and transgene copy number will vary for each founder animal. Some of the initial transgenic animals showed that the transgenes

were expressed in the expected fashion. They appeared to show the tissue specific and developmental pattern that was predicted for them (Brinster *et al* 1981, Chada *et al* 1985). However, this was not always the case and transgenes were often inappropriately expressed. This could manifest in a number of ways:

- ectopic expression whereby expression occurred in tissues other than the predicted ones
- incorrect timing of expression during development
- variation between lines, or individuals within a line
- variation within a tissue on a cell to cell basis

Expression of transgenes is influenced by a number of factors: those extrinsic to the transgene i.e. the structure and genomic position produced at the integration event, and those intrinsic to the design of the transgene i.e. enhancers (Clark *et al* 1993)

1.3.1 Extrinsic factors affecting transgene expression

Position effects are thought to be due to the influence of the integration site on the expression of the transgene. The nature of the flanking host sequences (i.e. the state of methylation, acetylation and chromatin structure; see section 1.7) is important for the expression of foreign DNA. It maybe that a region is more permissive for expression due to the absence of repressors or the presence of activators (Clark *et al* 1994). Position effects have been used to explain poor levels of expression, ectopic expression and variability within a transgenic line (Dobie *et al* 1997).

Al-Shawi *et al* (1990) described a position effect on transgene expression in mice. A transgene containing the major urinary protein (MUP) gene promoter and the Herpes simplex virus thymidine kinase (HSV TK) gene was used to generate five transgenic lines. In four of these high TK expression was found in the liver, testes and preputial gland. Only one did not have expression in the preputial gland. The integration locus of the aberrant line was cloned and used to generate secondary transgenic animals. Very little surrounding genomic DNA was included in the cloned transgene. All of the secondary transgenic lines expressed the transgene at high levels in the liver, testes and preputial gland. This demonstrated that the unusual expression was a consequence of the transgene's original chromosomal location and was not due to a mutation in the transgene.

Positive position effects can be utilised by using flanking chromosomal sequence isolated from actively expressing transgene loci, to enhance an alternative transgene's expression. A transgene containing the ovine BLG promoter sequence linked to α1-antitrypsin (AATB transgene) gave high levels of expression in the mammary gland. One line with very high levels of expression was investigated. The locus of the transgene was cloned into cosmids and was used to construct secondary transgenic animals. The cloned murine sequence conferred substantially improved expression upon the AATB transgene, which now showed position independent expression (Cranston *et al* 2001).

Alternatively, once a permissive genomic location is identified it can subsequently be used to facilitate expression for a number of constructs (Wallace *et al* 2000). In this work, the genome was scanned for "neutral" sites that were permissive for a number of transgenes. ES cell clones were

chosen for expression of a randomly integrated HPRT marker contiguous to an Oct4/lacZ transgene. The clones were assessed for appropriate regulation of the lacZ reporter upon differentiation. Clones that showed the correct regulation of lacZ also showed appropriate regulation of lacZ from three other developmental promoters that were subsequently targeted to the site.

The copy number of the array and the structure are also known to play a role in transgene expression. Some constructs are found to be copy number dependent i.e. the more copies present the higher the level of expression. In a number of cases however, transgenes in mice have been found to decrease their levels of transcription as the copy number of the transgenic array increases (Davis and MacDonald 1988, Sharpe *et al* 1992). Repeat sequences may also have a direct role in transgene silencing and this is discussed further in 1.4.

Modifier genes within the genome have also been shown to influence transgene expression. These strain specific genes can influence the expression of an endogenous gene or transgene (Nadeau 2001). Modifier genes have been shown to influence disease severity in humans and many examples of modifier genes exist in *Drosophila* (Birchler *et al* 2001). The expression of the modifier gene can have effects on a transgene at different levels. In one study a HSVtk lacZ transgene expressed at different levels depending on the genetic background of the mice (Allen *et al* 1990). If the transgene was placed in a DBA/2 or 129 background its expression was enhanced, whereas if it was placed in a BALB/c background its expression was decreased. The differences in expression correlated with differences in the methylation of the transgene, whereby the transgene was progressively

silenced in the BALB/c background. The modifier genes from the BALB/c background were found to be dominant suppressors of the transgene expression.

1.3.2 Intrinsic elements affecting transgene expression

Factors intrinsic to the transgene also affect expression. The locus control region (LCR) is defined as a DNA segment that confers transgene expression that is independent of its integration site in the host genome, but dependent on its copy number. A locus control region (LCR) responsible for position independence was identified in the β-globin cluster located ~30kb upstream of the first gene and is characterised by erythroid specific DNAse I hypersensitive sites (Ellis *et al* 1996). Similar LCR elements have been identified in the 3′ flanking sequence of the CD2 gene (Festenstein *et al* 1996). It is thought that the LCRs influence the chromosomal structure around a transgene allowing it to overcome position effects. In the case of the β-globin they also act as an enhancer of globin gene expression (Caterina *et al* 1994). The incorporation of elements that overcome chromosomal position effects is an obvious strategy for improving expression from transgenes. Both the LCR from β globin and CD2 have been shown to confer position independent expression on heterologous genes, although the effect is seen only in the expected cell type. However, this approach does not always guarantee position independent expression on an unrelated but contiguous transgene (Guy *et al* 1997). Another drawback to this approach is that many genes do not have a defined LCR.

MARs (matrix attachment regions) are defined as DNA sequences that can interact with the nuclear matrix and are also implicated in regulation of gene expression (Cockerill and Garrard 1986). MARs can act in a number of ways in a transgenic situation. They can disrupt the interaction between a promoter and enhancer leading to poor expression (Stief *et al* 1989). MARs have also been shown to insulate a transgene against position effects by acting as insulators (Namciu *et al* 1998). It is thought that MARs naturally act to define chromatin boundaries, and by doing so shield genes within one gene domain from regulatory effects from elements in another domain (Bonifer *et al* 1997).

A lack of introns has been shown to affect expression in a number of systems. Introns may affect expression because they contain regulatory elements required for expression. Both positive and negative regulatory elements have been discovered in introns. (Franklin *et al* 1991, Brooks *et al* 1994). Generic introns have also been found to increase gene expression when used in transgenic mice (Choi *et al* 1991). Introns may also act in a structural sense allowing transgenes to be processed as an endogenous gene would be. The use of genomic sequence may therefore help increase expression.

Gene loci of higher organisms have complex structural features. Eukaryotic genes are regulated by a number of elements that can be spread over kilobases surrounding the coding region. The variation of transgenic expression may be due to the fact that transgenes lack the necessary distal elements. By introducing genes with large flanking regions this may help improve efficiency without needing to identify the elements involved. The development of P1 vectors, bacterial artificial chromosome (BACs) and yeast

artifical chromosomes (YACs) allow much larger DNA inserts to be manipulated *in vitro*. P1 vectors can take fragments up to 80kb in size, while YACs can have inserts several hundred kilobases in length (Clark *et al* 1994).

The mouse tyrosinase gene has been successfully expressed as a YAC transgene (Schedl *et al* 1993). In this experiment the 250kb YAC contained 80kb of the tyrosinase coding region and 155kb of upstream sequence. The YAC was microinjected into an albino mouse strain allowing expression to be determined by the amount of pigmentation in the skin and eyes of the offspring. The founders and their offspring exhibited a homogenous strong agouti coat colour, in contrast with previous mini-tryosinase transgenes, which had shown variegated expression. The use of YAC vectors may also allow the introduction of very large genes or linked gene families (Lamb and Gearhart 1995).

1.3.3 Elements that interfere with expression

Prokaryotic vector DNA has been shown to inhibit transgene expression if it is not removed before microinjection of the DNA (Townes *et al* 1985, Chada *et al* 1985). Prokaryotic sequence may have different GC nucleotide content, which may indicate to the cell that it is foreign DNA and be subsequently silenced as a defence mechanism for the cell.

Elements such as mammalian cDNA or prokaryotic reporter sequence can silence transgenes (Clark *et al* 1997, Cohen-Tannoudji *et al* 2000). CAT and lacZ are often used as a reporter gene for transgenic studies but in some examples these can be silenced and can also silence associated transgenes.

1.4 SILENCING EFFECTS

As discussed in part 1.3.2, it is important to ensure that all the necessary elements are included in the design of a transgene to try and maximize its expression. However, in many cases the inclusion of all known regulatory elements is not enough to avoid silencing effects. These silencing effects can range from complete silencing of the transgene, variation within lines, to stochastic silencing of the transgene from cell to cell within the tissue of interest. It is therefore clear, that the lack of transgene expression can be considered from two different perspectives: either a failure to express due to intrinsic structure of the transgene or as an active silencing by the genome. If these two perspectives are combined, the lack of transgene expression can in many cases be thought of as the failure of the transgene to overcome the active silencing of the genome. In this section the idea of active silencing is reviewed with respect to transgene silencing prevalent within a number of systems from plants to mammals. The abundance of silencing effects found in the wide variety of organisms may indicate common mechanisms are employed throughout.

1.4.1 Silencing in Plants and Fungi

The drive to modify commercial crops using transgenic plants uncovered a plethora of gene silencing (Gallie 1998). This section will focus on silencing due to repeated transgenic sequences. When transgenic approaches were first used, more copies of a gene were expected to result in higher expression. Homology-dependent gene silencing (HDGS) in plants threw this notion on its head. Two types of HDGS have now been distinguished depending on the level of silencing. Transcriptional gene

silencing (TGS) is characterised by altered methylation and chromatin structure, while post-transcriptional (PTGS) involves RNA turnover aberrations (Kooter *et al* 1999).

TGS in plants involves the methylation and condensation of the allele to be silenced. Currently there are numerous examples in plants where the insertion of multiple copies of a transgene leads to loss of expression in some or all copies of the transgene. Repeat induced gene silencing (RIGS) in *Arabidopsis* was found to be dependent on repeated sequence, was reversible and correlated with decreased steady state mRNA and increased DNA methylation. The authors propose that it reflects ectopic interactions of paired homologous DNA with flanking heterologous DNA, inducing condensation into a non-transcribable state by the exchange of chromatin components (Assaad *et al* 1993).

Alternative theories suggested that transgenes may physically interact or compete for non-diffusible essential factors i.e. the nuclear matrix. Alternatively, specific RNA degradation may occur due to either antisense RNA production or accumulation of RNA triggering degradation (Flavell 1994, Baulcombe 1996, Matzke *et al* 1996). Recently evidence suggests homology dependent gene silencing in *Arabidopsis* is not occurring through an RNA mechanism (Luff *et al* 1999). A promoterless inverted repeat construct produced homology-based methylation with no RNA from the transgene being found. This indicated that the inverted repeat structure itself was responsible and not an RNA or protein product. The authors suggest the methylation is induced by DNA-DNA pairing, possibly involving the formation of hairpin structures, which could be targeted for methylation. The data available at this time suggests that methylation is

necessary for silencing at some loci in *Arabidopsis*. An *Arabidopsis ddm1* DNA hypomethylation mutant reduced the levels of methylation on a silenced gene (Jeddeloh *et al* 1998). The progressive loss of methylation was correlated with a loss of gene silencing.

Co-suppression involves reciprocal silencing i.e. all copies are silenced, including the transgene and endogenous homologs (Flavell 1994, Matzke and Matzke 1995). Experiments showed that critical regions of homology for silencing could be the promoter, or the coding region or both. Repeated sequences that were silenced could be located *in cis* or on different homologous chromosomes. This appears to be a case of post-transcriptional gene silencing and is associated with enhanced RNA turnover in the cytoplasm (Matzke and Matzke 1998).

In PTGS, RNA with sequence homology to the introduced DNA is found at lower levels than expected. It has been suggested that this sequence-specific effect at the post-transcription level is due to the expression of antisense RNA from the transgenes. The antisense RNA then duplexes with the target RNA, promoting its degradation or interfering with its translation. The theory of aberrant RNA expression or antisense RNA being expressed in cases of PTGS was supported by the finding that small antisense RNAs complemetary to the target mRNA have been found in PTGS plants (Hamilton and Baulcombe 1999). These 25 nucleotide antisense RNAs were never found in the absence of silencing, and possibly spread throughout the plant, thus explaining the previously reported systematic spread of PTGS from one infected leaf to the whole plant (Voinnet *et al* 1998). It was suggested that these RNA are transcribed from an RNA template. This would require the presence of an RNA-directed polymerase (RdRP). Indeed

a RdRP has since been cloned in plants (Schiebel *et al* 1998). Lending weight to this theory is the finding that PTGS mutants in *Neuorspora* encode for a protein showing homology to the plant RdRP (Cogoni and Macino 1999).

Paramutation in Maize plants occurs when a paramutable allele is converted to a new state of activity displayed by another silencing allele (paramutagenic), which may revert if the alleles segregate (Matzke and Matzke 1998). Other trans-silencing effects in transgenic plants have similarities with this phenomenon. For example, some methylated trans-silencing loci can transcriptionally inactivate an unlinked loci that they share homology to, and methylate it. The target loses methylation only gradually after segregating from the silencing locus (Jakowitsch *et al* 1999).

Fungi also have a number of silencing effects which share some similarities to the processes found in plants (Irelan and Selker 1996). These processes avoid the duplication of repetitive sequences by inactivating or mutating them. Pairing of repeat sequences is implicated in RIP (Repeat induced point mutation) and MIP (Methylation induced premeiotically). In *Neurospora crassa* RIP modifies duplicated sequence by C - T transititions in a process often accompanied by frequent methylation, inhibiting transcription. In *Ascobolus*, the process of MIP shows repeat sequences become methylated but are not mutated. The methylation is found to be exactly co-extensive with the duplicated sequence, supporting the idea that DNA pairing is responsible (Matzke and Matzke 1995).

All of the systems in plants that are involved in silencing repeats are thought to be part of a genome defence mechanism against invasive DNA sequences and viruses (Ratcliff *et al* 1997, Matzke and Matzke 1998). The

idea is that PTGS would defend against viruses with an RNA genome, while TGS would defend against transposable DNA elements (Kooter *et al* 1999). It has now been found that PTGS is suppressed by virus-encoding proteins which also suppress the anti-viral defense in the host (Anandalakshmi *et al* 1998).

1.4.2 Silencing in transgenic *Drosophila*

As with plants there are many examples in *Drosophila* where transgenes do not express as expected (Dorer 1997). Often the repeated nature of the transgene appears to be involved in the silencing.

Closely linked repeats of a *Drosophila* P transposon carrying a *white* eye pigment transgene were shown to cause *white* variegation. More than three copies of the transgene in an array showed a variegating mutant phenotype that strengthened with increased copy number. The degree of variegation was affected by the relative distance of the transgene to constitutive heterochromatin (Dorer and Henikoff 1994). The transgenes were also responsive to protein components of heterochromatin (Powers and Eissenberg 1993). The authors suggested that pairing of these repeat sequences at a single locus may lead to the formation of chromatin structures that are preferentially recognised by heterochromatic specific proteins i.e. HP1 (James *et al* 1989, Eissenberg 1990, Belyaeva *et al* 1993). These heterochromatic proteins would in turn exclude the transcriptional machinery from the locus (Dorer and Henikoff 1994).

Sabl and Henikoff later presented a case where X-ray induced chromosomal rearrangement placed natural heterochromatin next to the

genomic sequence of the *brown* eye pigment gene, which was subsequently manipulated with P transposase. This system allowed them to remove or increase the copy number (Sabl and Henikoff 1996).

It was found that as the copy number increased a decrease in the frequency of expressing cells occurred, indicating that the variegation had increased. The copy number and orientation of the transgene determined the phenotype in this system. This observation paralleled those showing variegation increasing with increasing copy number described above in mini *white* transgenes (Dorer and Henikoff 1994, Sabl and Henikoff 1996). These effects were unlikely to involve any feature specific to both *brown* and *white* genes, indicating heterochromatic silencing is a general property of repeated arrays. HP1 containing complexes might specifically bind somatic paired structures that form at locally repetitive structures. Mini-*white* repeat arrays were later shown to be able to cause silencing of a vital gene close to the array *in cis*. This spreading of heterochromatic silencing correlated with the array size and the repeat array behaved like natural heterochromatin (Dorer and Henikoff 1997).

More evidence for the involvement of heterochromatin in repressing repeated transgenes in *Drosophila* is presented by the selective accumulation of Polycomb group proteins at repressed loci (Pal-Bhadra *et al* 1997). Mutations in the Pc-G gene resulted in a 50% reduction in co-suppression. The recruitment of these proteins to repeated transgenes may help localise the DNA sequence to an inactive nuclear compartment i.e. the chromocentre of *Drosophila* where a number of heterochromatic domains are located together.

1.4.3 Silencing in transgenic mice

Many examples of variegation in mouse transgene expression have been documented over the years. One early paper found that a *tk* gene integrated in mouse centromeric heterochromatin was affected by DNA methylation and *cis* effects from the flanking region. The repression of *tk* was thought to reflect position effect variegation (Butner and Lo 1986). In other cases, as discussed in 1.3.1, strain-specific modifiers influenced transgene expression (Allen *et al* 1990). Other papers documented variegation in coat colour from tryosinase transgenes (Bradl *et al* 1991, Mintz and Bradl 1991).

1.4.4 Repeat Induced Gene Silencing in transgenic mice

Recently data from mammalian systems have confirmed that repeat induced gene silencing is a widespread phenomenon (Wolffe 1997). Within certain high copy number transgenic lines, animals would exhibit variegated expression within a tissue (Dobie *et al* 1996). It was predicted that transgenes in animals would form heterochromatin and be silenced in a similar manner to repeats found in plants or *Drosophila* (Dorer 1997). Since the mid 1990s it has now become much clearer that the repeated nature of multicopy arrays may be as influential a cause of variegation in transgenic mice as position effects are.

In some cases a correlation between the increasing copy number, and the increasing methylation of the transgenes promoter has been observed (Mehtali *et al* 1990). Methylation of DNA in animals is known to be involved in repression of transcription. In other cases, the variegated levels of

expression did not correlate with an increase in the methylation of the transgene, but did correlate with a restricted chromatin structure noted as a loss of DNAse hypersensitive sites (Garrick *et al* 1996). Both these epigenetic factors may have a crucial role in silencing repeats.

Robertson *et al* (1995) produced transgenic lines expressing β-galactosidase controlled by globin promoters linked to the major tissue specific regulatory element of the α-globin locus. This system permits the analysis of transgene expression in individual red blood cells. The analysis showed expression of transgenes within lines is heterocellular and suggests the variation in total transgene activity between lines is mainly due to the number of cells committed to transgene expression rather than different levels of promoter activity.

As discussed in 1.4.1 and 1.4.2 Repeat Induced Gene Silencing (RIGS) was shown to decrease the proportion of cells that expressed the transgene in systems such as *Drosophila* and plants. In mice however, due to the use of microinjection for generation of transgenic lines, studies could not distinguish between position effects and the number of copies integrated. In 1998, after this PhD project was started, it was formally established that the tandem repetition of transgenes could produce transcriptional silencing in mammals (Garrick *et al* 1998). High copy number lines were established by microinjection. The transgene used was a hybrid lacZ reporter construct driven by the human α-globin promoter and the αHS40 enhancer-like element, with a loxP site integrated. The loxP site did not interfere with the transgene expression. Using Cre/lox recombination transgene copy number was then altered at the same integration site. Expression levels were

analysed for the parental unreduced line and for the reduced line in 12.5 dpc embryos using lacZ staining. Two high copy lines (>100 transgene copies) showed a variegated pattern of expression with less than 1% of erythroid cells containing an active locus. Reduction in copy number to five or one copy resulted in a large increase in the transgene expression level. The reduction in copies correlated to suppression of variegation. Expression had increased over 1000-fold in the line reduced to five copies, while there was a 180-fold increase in the other line reduced to one copy. This expression increase was correlated with a decrease in chromatin compaction and methylation at the transgene locus. Run-on analysis of the transgene indicated that the silencing in the unreduced line was at a transcriptional level.

By contrast, other studies into RIGS have failed to find an increase in the proportion of cells expressing after reduction in copy number (Ramirez *et al* 2001). A keratin5-lacZ construct was shown to express mosaically. The transgene expression was studied in the basal layer of stratified epithelia. Six K5Zlox lines were generated; four of which variegated and two were silenced. Using Cre/loxP the transgene copy number was modified either by microinjection of Cre, or by breeding to a line of mice expressing Cre specifically in the stratified epithelia. Tissue sections of one parental line (14 copies) showed no difference to the reduced line (1-2 copy). Another line showed a reduced intensity of β-gal staining, but not in the proportion of cells expressing. These results indicated that copy number reduction had little effect on variegation of K5Z transgenes. This discrepancy between this work and that of Garrick was explained by the authors as differences

between the two systems studied; i.e. the cell types, the transgene structure or that different chromosomal sites may be more prone to repression.

In all the situations discussed, there has been the suggestion that the genome has mechanisms that can identify foreign sequences, and can then silence them as a defence mechanism to maintain the gene expression pattern of that cell. How cells recognise foreign DNA is not clear, although the difference in codon usage, the GC content of a transgene relative to its chromosomal location or the repeated nature of inserts all may trigger the many diverse silencing mechanisms that exist. Bacterially derived sequence of the lacI repressor can be eliminated from somatic cells (Scrable and Stambrook 1999). However, the same sequence recoded for mammalian codon usage was expressed and was hypomethylated. Therefore, cells do seem to possess an ability to distinguish mammalian-like sequences from foreign sequence.

The genome of vertebrates is made up of mosaics of isochores, which are very long stretches (>300kb) of DNA that are homogeneous in base composition and are compositionally correlated with the coding sequences that they embed (Bernardi 1995). Isochores can be partitioned in a small number of families that cover a range of GC levels. Insertion of transgenic DNA into isochores that differ in their GC content may be involved in helping cells to recognise foreign DNA. It has been found that there are tight constraints on base substitutions in GC rich isochors even if the region is non-coding, indicating these sequences must be functionally important. The GC richness of the foreign sequence would alert the surveillance system, the

intruding DNA would become methylated and then in some cases removed, from the genome.

The examples from *Drosophila,* mice and plants show how the repeated nature of transgene arrays plays an important role in transgene expression. It is argued that arrays in transgenic animals are subjected to similar mechanisms of silencing as discussed, because of their repeated nature. However these effects may still be dominated by the chromosomal context that the repeats find themselves in.

1.5 POSITION EFFECT VARIEGATION

Given the stochastic nature of silencing in many transgenic systems comparisons have been drawn to position effect variegation, described in *Drosophila* in 1930 (Muller 1930). In these early classic experiments chromosomal rearrangements were induced by X-ray radiation and shown to lead to variegated expression of a gene. Thus, the translocation of a euchromatic gene close to a region of heterochromatin was thought to subject it to the spread of the heterochromatin and consequently silencing (Henikoff 1990). This was thought to result in mosaic expression of translocated genes on a cell-to-cell basis depending on the extent of heterochromatization.

PEV has been extensively studied in *Drosophila* leading to the discovery of genes suppressing or enhancing variegation, as well as implicating heterochromatic proteins in this phenomenon. Some suppressor mutations including Suppressor of variegation *Su(var)205* and *Su(var)3-7* encode for HP1 and a zinc finger protein, both of which are heterochromatin

components, as predicted for genes that reduce silencing as their dose is reduced (Eissenberg *et al* 1990). Several enhancers encode transcription factors, while one enhancer locus, D-Ubp-64E, encodes an ubiquitin-specific protease (Henchoz *et al* 1996).

Polycomb proteins (Pc) have been found to be closely associated with PEV. These proteins are involved in *Drosophila* development. They act to repress transcriptional activity by controlling the organisation of entire chromosomal domains. Polycomb silencing can spread along the chromatin fibre from the Polycomb response elements (Birchler *et al* 2000). It has also been suggested that Pc can spread *in cis* and *in trans* to produce co-suppression of non-homologous transgenes (Pal-Bhadra *et al* 1999).

The properties of PEV discussed above, help highlight the similarities to transgene silencing. In both cases, the degree of variegation is affected by the relative distance of the gene to constitutive heterochromatin (Dorer and Henikoff 1994, Dobie *et al* 1996). The influence of heterochromatic proteins i.e. HP1, is shared between PEV and transgene silencing (Eissenberg *et al* 1990, Festenstein *et al* 1999). The stochastic manner in which PEV affects the gene expression also draws comparisons with transgenic examples (Henikoff 1990, Robertson *et al* 1995).

1.6 EPIGENETIC EFFECTS

Position effect variegation and repeat effects both fall in to a wider area of study, that of epigenetics. This field concerns itself with a wide number of phenomena which all affect gene expression at a level other than the DNA sequence (Gasser *et al* 1998). Common mechanisms may exist

between many diverse effects. Methylation, chromatin, histone acetylation, nuclear architecture and localisation all play various roles in the epigenetic regulation of gene expression (see 1.7). Given the need for tight control of gene expression throughout development and the various tissues of the body, it is not surprising that such a complex web of factors controlling gene expression has arisen. Imprinting and X inactivation are two of these effects discussed below. These are interesting epigenetic phenomena as both show some similarities to PEV and repeat effects; mainly 'counting' and the involvement of methylation.

1.6.1 Imprinting

Genomic imprinting is involved in gene regulation by ensuring imprinted genes express only one of the two parental copies in a parent-of-origin specific manner (for review of imprinting, see Reik and Walter 2001). It provides a puzzling exception to Mendelian laws of inheritance as both parental sequences encode the same gene. A mark or imprint conferring memory must exist that differentiates between the parental genomes, be inherited during somatic cell division but is erased in the germline (Falls *et al* 1999).

One theory suggests imprinting arose due to parent–offspring conflict in the amount of maternal resources allocated to an embryo (Moore and Haig 1991). This theory predicts that imprinted genes expressed from the paternal genome should enhance growth, while maternally expressed genes should inhibit growth, providing a conflict between the two parental genomes. Indeed many of the genes found to be imprinted are involved in

fetal growth, placental development or behaviour designed to gain maternal attention (Mochizuki *et al* 1996).

Imprinted genes have been shown to be differentially methylated depending on the parent-of-origin. These differentially methylated regions or domains (DMR or DMD) are thought to control the expression of imprinted genes (Reik and Walter 2001).

Mouse *H19* and *Ifg2* show inverse patterns of expression. The mouse *H19* gene is maternally expressed, whereas *Igf2* is paternally expressed which fits with the conflict theory of imprinting (Thorvaldsen *et al* 1998). The paternal-specific methylation upstream of *H19* appears to be an epigenetic mark required for imprinting and leads to the paternal suppression of *H19*. The deletion of the paternal H19 DMD activates expression in the normally repressed paternal *H19*. It also causes reduced *H19* expression in maternal inherited transgenic mice, along with activation of *Igf2*. The DMD is therefore required on both parental alleles for reciprocal imprinting of *H19* and *Igf2* (Thorvaldsen *et al* 1998). It was later shown that this DMD contained an element that blocked enhancer function (Bell and Felsenfield 2000, Hark *et al* 2000). The enhancer blocking activity was found to be dependent on the protein CTCF. If the binding sites for CTCF were methylated, the enhancer blocking activity was lost.

Imprinting at this locus is controlled by methylation of the paternal allele, which inhibits CTCF binding and the insulator element activity. From this natural situation it is clear that methylation plays an important role in imprinting. The control of imprinted gene expression by methylation reflects the situation found with transgene silencing where many examples

exist whereby the transgene silencing is accompanied by an increase in methylation (Allen *et al* 1990, Mehtali *et al* 1990).

1.6.2 X inactivation

In mammals, due to heteromorphic sex chromosomes a mechanism is required for dosage compensation between females (XX) and males (XY). X inactivation occurs early in mammalian female development. This is achieved by global silencing of one X chromosome. This process requires that the cell counts the chromosome number, chooses which X chromosome to silence, initiates and then maintains inactivation (Marahrens 1999). Once inactive, the X chromosome becomes heterochromatin like and is condensed into a Barr body (Lyon 1998).

In diploid cells one X chromosome remains active, regardless of how many X chromosomes exist in the cell. The choice of which X to inactivate is a random decision made by individual female cells. However, the Xce (X-controlling element) can influence this decision. Once the decision has been taken, the inactive state is stably maintained in somatic cells. An X-inactivation centre (Xic) was discovered which is required for the initiation of silencing but not for the maintance of inactivation (Brown and Willard 1994). The *Xist* gene encodes a 15kb untranslated mRNA and maps to the Xic. This mRNA is expressed exclusively from the inactive X, is induced prior to onset of inactivation and is found to 'coat' the inactive X (Brown *et al* 1991). *Tsix* is a 40kb RNA that starts 15kb downstream of *Xist* and is transcribed across the Xist locus (Lee *et al* 1999). Before X inactivation, *Tsix* is expressed from both X chromosomes. At the onset of inactivation *Tsix* becomes monoallelic and is found to associate only with the future active X. It then persists until X

inactivation is established. Its antisense nature suggests the possibility that *Tsix* directly blocks *Xist* action on the X chromosome. *Tsix* is now thought to regulate Xist *in cis* and determines which chromosome is silenced without affecting the actual silencing (Lee and Lu 1999).

The initiation of heterochromatin formation at X inactivation has been proposed to occur due to pairing of the homologous X chromosomes (Marahrens 1999). This hypothesis is supported by the fact that *Xist* transgenes will only inactivate autosomal sequences if they are integrated as tandem arrays, which fits with a model of pairing reactions between adjacent transgene copies (Heard *et al* 1999). Repetitive LINE elements on the X chromosome may be acting as boosters of the spread of Xist RNA. The RNA is thought to help bring the LINE sequences together, inducing a 'repeat response' from the cell, which heterochromatizes the X chromosome (Lyon 1998).

X inactivation is a natural example of gene dosage, which appears to involve counting and homology induced silencing. This has strong similarities to transgene silencing, which has been shown in many diverse systems (plants, *Drosophila* and mice) to involve homology based silencing effects (Wu and Morris 1999). X inactivation leads, as do many examples of transgene silencing to the heterochromatinization of the region (Assaad *et al* 1993, Dorer and Henikoff 1994, Garrick *et al* 1998). The ability to 'count' either imprinted genes or X chromosomes and the systems to inhibit expression from them, i.e. methylation, pairing and heterochromatinization, show how transgene silencing may use silencing systems that already exist within the cell.

1.6.3 Epigenetic inheritance

Genetically identical individuals show phenotypic variation, which has previously been described as 'developmental noise' or as variable penetrance due to multi-trait modifiers. Breeding strategies to discount genotypic and environmental differences have still shown variation in both endogenous and transgenic expression (Morgan *et al* 1999, Sutherland *et al* 2000). Evidence is accumulating to suggest that this phenotypic variation is due to epigenetic modifications of the genome being inherited (Rakyan *et al* 2001).

These epigenetic modifications which include methylation, chromatin and histone acetylation, function to affect gene expression, which has a knock-on effect on the phenotype of an organism. It was previously thought that all epigenetic marks were erased and reset on passage through the germline. Recently, evidence has started to emerge that suggests that sometimes these epigenetic marks may persist, and be inherited in the next generation (Morgan *et al* 1999). For example, inbred mice from over 20 generations would be expected to be phenotypically indistinguishable. However, mice carrying the agouti yellow viable allele (A^{yv}), show varying colour patterns from full agouti to full yellow. A trans-generational effect was present, coat colour of grandmother related to the coat colour of the second generation. However, due to the genetic background, genetic modifiers could be ruled out. Uterine environment was also excluded as a source of the variation. This leaves the inheritance of an epigenetic mark from generation to generation as the most plausible explanation.

From the examples described in section 1.6, it is clear that in some cases transgene silencing shows similarities to the natural cases of gene silencing found in the genome, i.e. imprinting and X-inactivation. It may be that transgene expression is controlled by the existing systems of gene regulation within mammalian cells.

1.7 MECHANISMS OF EPIGENETIC REGULATION

Epigenetic control of gene expression is a complex subject. Many diverse phenomena may have related mechanisms or involve the same key factors. In this section some of the 'biochemical' mechanisms involved in epigenetic regulation, which may have similar roles between a wide range of organisms, are discussed.

1.7.1 Methylation

Cytosine methylation of CpG dinucleotides has been suggested as a host defence against intragenomic parasites (Yoder *et al* 1997). Active transposons threaten the regulated expression of the genome in a number of ways. Insertional mutagenesis, recombination between transposons causing translocations and rearrangement, aberrant RNA expression from a transposon or antisense RNA to an endogenous gene could all disrupt cellular expression.

It has been suggested that high numbers of spontaneous mutations in *Drosophila* are due to transposable elements. It must be considered at this point that *Drosophila melanogaster* has no detectable endogenous methylation (Urieli-Shoval *et al* 1982). *Drosophila,* having no methylation appears unable

to defend its genome against these intragenomic parasites. The expression of mammalian methyltransferases causes genomic methylation and lethality in *Drosophila*. This lethality of genomic methylation suggests that methylation has functional consequences for *Drosophila* development (Lyko *et al* 1999). This data suggests *Drosophila* at some point had genomic methylation and has since lost it. It has been shown directly that transposable elements can be silenced by HDGS, providing a link to the mechanisms that silence transgenes and a 'defence' system (Jensen *et al* 1999).

Methylation is necessary for proper embryonic development, as mice lacking a functional DNA methyltransferase fail to develop properly and die in mid-gestation (Li *et al* 1992). Methylation has been suggested as a mechanism to control inappropriate gene expression from a complex genome in a global manner. It is clear that it is just as important to selectively control the genes to be switched off, as it is to control the expression of genes through development. Given the number of genes active in a differentiated cell compared to the overall gene number, repression of genes must be a crucial role in gene control (Bird 1995, Siegfried and Cedar 1997).

Methyl-CpG-binding proteins MeCP1 and MeCP2 interact specifically with methylated DNA and mediate transcriptional repression (Boyes and Bird 1991, Nan *et al* 1998). MeCP2 binds to chromosomes in a methylation dependent manner, and contains a transcriptional-repression domain (TRD) that can function *in vitro* and *in vivo*. It has subsequently been shown that a region of MeCP2 associates with a co-repressor complex including mSin3a and histone deacetylases (HDACs). This work showed that MeCP2 provides a bridge linking methylation to histone deacetylation and subsequently repression (Nan *et al* 1998).

More links between methylation, acetylation and chromatin have recently been found. Dnmt1 (DNA methyltransferase 1) is found to be associated with histone deacetylase activity *in vivo* (Fuks *et al* 2000). HDAC1 can bind to Dnmt1. A transcriptional repression domain in Dnmt1 functions by interacting with the catalytic domain of HDAC. Two models exist for the transcription repression of these proteins: the Dnmt1 associated deacetylase activity may be required to deacetylate newly formed nucleosomes, or the deacetylase activity may be needed first to remodel the chromatin to allow methylation to take place.

The pattern of overall methylation involves the methylation and demethylation of DNA throughout development. Recently a demethylase was identified (Bhattacharya *et al* 1999). The protein is a member of a conserved family of Methyl-CpG-binding domain (MBD) proteins, and shows homology to MeCP2 at the MBD. The enzyme was thought to convert 5-methylcytosine to cytosine and methanol. This demethylase is identical to MBD2b. However, the *in vivo* function of this demethylase has not yet been proven (Wolffe, Jones and Wade 1999).

Currently four mammalian DNA cytosine methyltransferases (DNMTs), five methyl-CpG-binding proteins and a demethylase have been found (Bird and Wolffe 1999). An alternative role for methylation is that the default setting of the genome is repression, and that demethylation is the key to expression. Investigations into the complex connections between methylation and expression are continuing. Nevertheless it is quite clear that methylation is an extremely important factor regulating endogenous and transgenic gene expression in mammals. The important role of methylation

is shown by the correlation of aberrant methylation with cancer and some other human diseases (McBurney 1999, Hansen *et al* 1999, Tycko 2000).

Given the crucial role of methylation in controlling endogenous gene expression as discussed above, it is not surprising that there are numerous examples of transgene silencing involving methylation. Transgene sequences in mammals and plants, have been found to be methylated when silenced (Palmiter, Chen and Brinster 1982, Garrick *et al* 1998, Assaad *et al* 1993). When repeated transgenes were reduced to single copy and showed expression, there was found to be a correlated change in the methylation status (Garrick *et al* 1998). It may be that the methylation status of transgenes is crucial to their ability to express, similar to the control of endogenous gene expression (Thorvaldsen *et al* 1998).

1.7.2 Histone acetylation

The eukaryotic genome is maintained in a DNA protein complex called chromatin. The main protein components of chromatin are the core histones. Histones associate to form a nucleosome, where 146bp of DNA is wrapped around an octomer of histones. Histones are post-transcriptionally modified in a number of ways including phosphorylation, ADP-ribosylation, methylation and acetylation (Brownell and Allis 1996). Histone acetylation results in the neutralisation of a single positive charge. Acetylated histones are associated with gene expression, while deacetylation associates with a repressed state. This is due to the fact that histone acetylation weakens the association of histones with DNA, allowing a more 'open' nucleosomal structure to occur (Struhl 1998). Under-acetylation of histone H4 has been

found to be a general property of heterochromatin and is thought to play a role in the stabilisation of the inactive X chromosome (Keohane *et al* 1998).

Histone acetylases have been found to be closely associated with PolII machinery. p300/CBP histone acetylase is found in a wide range of organisms and is found to be tightly associated with the PolII machinery (Nakajima *et al* 1997). This provides a simple mechanism linking acetylation and transcription. HDAC have been found in large multi-protein complexes that include repressors (Struhl 1998). The targeting of histone acetylases and deacetylases to promoters most likely causes local modifications of chromatin structure.

Histone acetylation has also been proposed as an epigenetic mark that could survive from one cell generation to the next (Turner 1998). The model predicts that a level of information could be encoded in the pattern of lysine specific acetylation on the nucleosome surface, and that this could be turned into structural and functional effects by non-histone proteins. If the acetylation can also locate an appropriate lysine specific acetyltransferase then this mark could be maintained through DNA replication.

Histone acetylation provides a link between methylation and chromatin remodelling (Bestor 1998). Transgenes can often be condensed into heterochromatin (Dorer and Henikoff 1994, Garrick *et al* 1998) and it may be possible that histone acetylation has a role, not just in regulating endogenous gene expression, but also that of transgenes. However, at this time no studies have been carried out into the direct role of histone acetylation in transgene expression.

1.7.3 Chromatin remodelling

When chromatin was first discovered it was thought to function only as a compaction mechanism for packaging the genome into the nucleus. More recently, as effects of methylation and acetylation have been studied it has been established that modulations in chromatin have direct roles in regulating transcription.

An *Arabidopsis* gene *DDM1* (decrease in methylation) was found to be required for maintance of normal methylation patterns. This gene was found to encode for a SWI2/SNF2-like protein. This discovery showed chromatin remodelling was an important process for maintaining DNA methylation (Jeddeloh *et al* 1999). Members of the SWI2/SNF2 family are involved in various functions including transcriptional co-activation or co-repression, chromatin assembly and DNA repair.

The SWI/SNF complex is a 2 MDa multi-subunit DNA dependent ATPase that regulates gene expression through changes in chromatin structure. SWI/SNF can alter nucleosome structure and may remodel nucleosomes *in trans* or *in cis*, by displacement or by sliding of the octomer (Peterson and Workman 2000). It is now thought that gene specific activators recruit the remodelling SWI/SNF complex directly to genes. This 'activator model' is supported by evidence of direct interaction between SWI/SNF and transcriptional activators (Yudkovky *et al* 1999). SWI/SNF highlights the interaction between methylation, nucleosome structure and position and the remodelling of chromatin in relation to gene expression.

The chromatin formation surrounding transgenes can correlate with its expression profile (Whitelaw 2000). The precise localisation of nucleosomes and their remodelling may be crucial to the decision to express a gene (Kornberg and Lorch 1999). Therefore, the mechanisms by which chromatin is remodelled may help elucidate the higher order regulation by which transgenes are expressed.

1.7.4 Nuclear localisation

Functional domains within the nucleus are becoming apparent through FISH experiments. The compartmentalisation of a gene may be involved in its subsequent expression.

Heterochromatic protein 1 is associated with heterochromatin but is not a histone (Eissenberg and Elgin 2000). HP1 plays a role in regulating gene expression through heterochromatin, but it may also play a role in nuclear localisation for gene expression. HP1 is encoded by a locus identified as a dominant suppressor of PEV, *Su(var)*2-5 in *Drosophila*. HP1 was later shown to share a chromo domain motif with the homeotic gene silencer *Polycomb* (Pc) (Paro and Hogness 1991). Over-expression of HP1 leads to a modification in PEV, specifically the number of cells expressing changes in a dose-specific manner but also in a chromosomal-context-dependent manner (Festenstein *et al* 1999).

HP1 can interact with the lamin B receptor (LBR), which is an integral membrane protein of the nuclear envelope (Jones *et al* 2000). HP1 has also been shown to interact with *Ikaros*, a transcription factor in mice. *Ikaros* has been used as a marker for centromeric heterochromatin, and has shown that

lymphoid specific genes are localised in the nucleus according to their expression status (Brown *et al* 1997).

Transgenes have been shown to circumvent silencing by avoiding nuclear compartments in which expression would be suppressed (Francastel *et al* 1999). Nuclear localisation may therefore play an important role in cases of transgenic PEV, where the repeat structure is recognised by chromatin remodelling and heterochromatic proteins, which may then help localise the transgene to a nuclear compartment that helps maintain the silent state.

1.8 TRANSGENE SILENCING IN THE MAMMARY GLAND

The mammary gland is an unusual tissue; it is only fully developed in the adult female after pregnancy, and undergoes involution after suckling has stopped, where upon it resembles the virgin gland (Mepham 1987). The expression of milk protein genes is controlled so that milk production is timed to coincide with the birth of pups. The production of milk is regulated by a number of hormones, including progesterone, prolactin and oxytocin (Hennighausen and Robinson 1998). The potential for foreign gene expression in milk, has led to extensive use of transgenes that express in the mammary producing animals termed 'bio-reactors' (Houdebine 2000).

1.8.1 β-lactoglobulin

β-lactoglobulin (BLG) is the major protein in the whey of ruminants and is found in the milk of other animals including horses, dogs and dolphins. However, it is not found in humans or rodents (Pervaiz and Brew 1985). The ovine gene encoding BLG is 7379 nucleotides long and is single copy in ruminants (Harris *et al* 1988). BLG protein exists as a dimer of two identical subunits of 162 amino acids each containing five cysteine, of which four form intra-chain disulphide bridges. Each subunit has a molecular mass of 18000 (18K). Approximately 5% of poly (A) RNA in the mammary gland of lactating sheep codes for BLG transcribed as an 800 nucleotide RNA (Mercier *et al* 1985).

Mammary development, lactogenesis and involution are all controlled by a complex system involving hormones, growth factors and inter- and extra-cellular interactions. BLG was shown to have three sites for milk protein binding factor (later shown to be Stat 5). Mutation of these sites reduced expression of BLG transgenes, while mutation of all three resulted in loss of hormone response *in vitro* (Burdon *et al* 1994). Prolactin is predominantly synthesised in the pituitary of vertebrates and is one of the main hormone regulators of milk expression along with estrogen, growth hormone and glucocorticoids (Tucker 2000). Prolactin activates BLG transcription through the JAK/STAT pathway. JAK2 is recruited to the prolactin receptor, phosphorylating the Stat5 isoforms 5a and 5b. The phosphorylated Stat then bind to BLG and the gene is expressed (Hennighausen and Robertson 1998). In lactating mammary the BLG promoter is found within a nuclease (DNase I) hypersensitivity site. This HS

site covers the Stat binding sites and its appearance correlates with the formation of a hormonally induced transcription complex (Whitelaw 1996). These changes in DNase HS at this region are not dependent on Stat5 interaction at the binding sites (Whitelaw 2000).

1.8.2 Lipocalins

BLG belongs to a phylogenetically conserved group of more than 40 proteins called the lipocalins (Flower 1996). Ovine BLG shows similar organization of exons and introns as retinol binding protein, major urinary protein, apolipoprotein D and alpha-1-acid glycoprotein. BLG and retinol binding protein show very similar three-dimensional structures (Ali and Clark 1988). These lipocalins function in binding and transportation of hydrophobic ligands, BLG has been shown to bind fatty acids and retinol. It is still unclear, however, what the biological role of BLG is.

1.8.3 BLG transgenic mice

BLG genes from a number of species including ovine, caprine and bovine have now been used to produce transgenic mice (Simons *et al* 1987, Ibañez *et al* 1997, Gutiérrez-Adán *et al* 1999). BLG is expressed at the highest levels using the ovine construct (23mg/ml; five times the estimated BLG concentration in sheep), followed by bovine (3mg/ml) and then caprine (0.5mg/ml). It may be however, that these differences are due to the varying amount of 5′ and 3′ sequence used in the different constructs, or the number of experiments sampled.

Expression of the BLG transgene in mice had no deleterious effects and the only difference between milk composition from control mice and the transgenic mice was the presence of BLG. Initially as with many transgenic systems, BLG was thought to be appropriately expressed. BLG expression in transgenic mice was shown to be very similar to the endogenous β-casein in both temporal and spatial expression. Both expressed at low levels in virgin glands and then gradually increased in the first ten days of pregnancy. Beyond ten days there was high expression, which continued till after parturition. These levels grew from 5% at day 10 to 65-68% at parturition compared to the levels of mid-lactation. BLG transgenes were also restricted to expression in the secretory epithelial cells (Harris *et al* 1991).

1.8.4 Bio-reactors

A hybrid transgene consisting of the BLG promoter fused to an α1-antitrypsin mini gene was shown to be capable of driving expression in the mammary gland of transgenic mice (Archibald *et al* 1990). This opened up the possibility of using transgenic animals as bioreactors for the production of pharmaceutically important proteins. The use of animals as bioreactors has become an extensive area of transgenic research and has been helped by recent advances in nuclear transfer (Clark 1998, McCreath *et al* 2000).

1.8.5 Determining intrinsic factors required for BLG expression

Initial experiments with BLG transgenes helped define the required elements for expression. A BLG transgene containing 408bp 5′ flanking region was sufficient to direct efficient expression of BLG to the mammary

gland (Whitelaw *et al* 1992). The proximal 408bp region had some characteristics of an LCR: tissue specific DNAse hypersensitive sites, interactions with tissue specific and general nuclear proteins. However if the promoter was reduced below the 408bp region it showed highly unpredictable expression patterns with many lines being completely silenced (Whitelaw *et al* 1992).

The 5′ flanking region of BLG contains all the necessary sequence to direct appropriate expression in transgenic mice (Webster *et al* 1995). The region comprises a cluster of binding sites for a number of transcription factors, contained in a mammary specific DNase I hypersensitivity site, and in many ways it is similar to known enhancers (especially INFβ enhancer; James *et al* 2000). The large number of transcription factor binding sites, and the DNAseI hypersensitivity site in the BLG promoter has drawn comparisons to LCR (James *et al* 2000). Experiments to determine if multiplying the number of transcription factor binding sites could function as a 'synthetic LCR' showed that the multimerisation of the BLG promoter failed to enhance expression, and in fact suppressed it. The multimer BLG promoter had failed to create a mammary specific DNase I hypersensitivity site. It was hypothesised that the transcription elements become targets for general repression if they fail to form the appropriate complexes required for activation (James *et al* 2000).

1.8.6 BLG MAR sequence

Although the minimal promoter can drive expression in some cases, regions downstream of the BLG promoter are required for efficient expression, as shown by the failure of the promoter to express a reporter gene in the mammary gland (Shani *et al* 1992, Webster *et al* 1995). It has now been shown that a single MAR exists within the BLG gene domain (Whitelaw *et al* 2000). The MAR was found in the proximal 3' flanking region, -95bp to –985bp to the polyadenylation signal. BLG is a GC rich gene (>60%) but the MAR is AT rich (average 55%). Removal of this MAR resulted in a lower basal level of expression in transgenic mice. By nuclear run-on assays, this reduction was determined to be due to a reduction in transcription. The reduced transcription was not due to the transgene failing to be induced by hormones. It was suggested the reduction corresponded to reduced activity from each transgene copy or possibly to a mosaic array were only a few transgene are active at any one time.

1.8.7 BLG and variegation

Early papers showed that BLG expressed efficiently in a tissue-specific manner and appeared to be free from position effects. The bovine BLG transgene has been shown to be position dependent and had an inverse correlation between copy number and expression (Gutiérrez-Adán *et al* 1999). When later studies looked at the expression of the ovine BLG transgene at the cellular level, variegated or mosaic expression patterns were discovered,

suggesting that position effects did indeed play a significant role in BLG transgene expression (Dobie *et al* 1996).

Dobie *et al* (1996) studied expression levels in three BLG transgenic mice lines. Two lines, BLG 7 and 45 exhibited variable levels of transgene expression. In these lines low expressing animals showed discrete patches of cells expressing the transgene. The variegation at the cellular level was mirrored by variation in the expression levels of BLG in milk: there was wide spread between sibs of the same line. Line 45 expressed from 16 to 30mg/ml while line 7 had an eight-fold difference between the lowest expressor and the highest. BLG protein levels were found to be stable within individual mice in successive lactations. The copy number within the lines appeared stable and no evidence of transgene rearrangement was found. Both variegating lines had over 15 transgene copies in arrays while the stable line had only two. The overall transgene protein expression level in milk was deemed to reflect the proportion of epithelial cells in the mammary gland expressing BLG mRNA. Chromosomal *in situ* hybridisation showed that the transgenic arrays in both variegated lines were located close to the centromere. The relatively large contiguous patches of cells where the transgene was active or inactive was consistent with clonal expansion. It was thus suggested that epigenetic silencing of transcription occurs stochastically in individual progenitor cells, which then transmits to the daughter cells giving rise to the mosaic patchy expression.

As discussed for other transgenic silencing effects the example of mosaicism or silencing in BLG transgenic mice has parallels to PEV in *Drosophila*, for example, integration of the transgene close to the centromere

is implicated in both examples. These studies however, could not distinguish between effects from the site of integration and the structure of the array. Dorer and Henikoff (1994) illustrated the importance of an increase in copy number in producing transgenic PEV like mosaicism, while both variable lines in the Dobie study had larger numbers of copies compared to the stable line. The data from many transgenic systems argues that mosaic expression and silencing can be due to both the location and the nature of the transgene array.

1.9 AIMS AND APPROACHES

The work in this thesis follows on from the work of Dobie *et al* (1996). In particular it addresses the hypothesis that the silencing of BLG transgenes was directly linked to the copy numbers of the transgene arrays.

This project attempts to address the issue of the importance of transgenic array structure in position effects and mosaic expression while avoiding the influence of the integration site. The Cre recombinase system enables the manipulation of an array *in vivo* - to reduce the copy number whilst maintaining the same integration site. With this approach it is possible to determine the extent to which the repeated nature of BLG transgenes contributes to silencing.

The initial aim of this project was to produce high copy number BLG-loxP transgenic lines by microinjection. These lines were used to investigate the transgene expression profile. Lines that showed variegation were used in further experiments. The study used the Cre-loxP system to produce a

reduction in copy number in two different ways. Firstly, BLG-loxP animals were bred to BLG-Cre animals, (which expressed the Cre recombinase specifically in the mammary gland) so producing a mammary gland specific reduction. Secondly, BLG-loxP fertilised oocytes were microinjected with a PGK-Cre construct, which produced a reduction throughout the animal and in its germline, allowing 'reduced lines' to be established. The expression profile of reduced animals was analysed to determine if any change in expression had accompanied the changes in copy number generated by this approach. These two approaches were compared and contrasted, and many of the issues related to Cre mediated recombination were addressed. My results show that reducing a tandem array of BLG repeats does not prevent stochastic silencing in the mammary gland but may, indeed, accentuate it.

CHAPTER TWO

MATERIALS AND METHODS

2.1 TRANSGENIC MOUSE LINES

2.1.1 Lox mouse lines

Lox lines of mice contain a 7kb BLG-loxP construct derived from BLGΔDp (Whitelaw *et al* 1992) with a LoxP site inserted at the 3' end (see Chapter 3). The BLG-loxP construct was digested with *Xho* I (see 2.5.7) and gel purified (see 2.5.9) to remove vector sequences which are known to interfere with transgene expression. Pronuclear stage eggs were obtained from superovulated C57BL/6 x CBA F1 females mated to C57BL/6 x CBA stud males. Miss Roberta Wallace microinjected the construct into either pronucleus of fertilised eggs at 1.5ng/μl, the cells were cultured, and cleaved embryos (2 cell stage) were transferred to pseudopregnant recipients (females that have been mated to vasectomised males), producing G0 pups. PCR and Southern blotting were used to identify positive transgenic pups. A transgenic line is defined as the transgenic offspring from a founder animal. Five Lox transgenic lines were established from these microinjections and were maintained by systematic crossing to C57BL/6 x CBA F1 hybrid mice.

2.1.2 BLG-Cre 74 mice

The generation of this line is described by Selbert *et al* (1998). Cre 74 mice contain a 6.7kb Cre fusion transgene driven by the BLG promoter. The Cre recombinase gene was cloned as a 2.5kb *Xho* I fragment into the *Eco*RV site of 4.2kb BLG/SK+ plasmid. The plasmid was linerised by *Not* I prior to microinjection into fertilised CBA x C57BL/6 eggs. This line was maintained by systemic crossing to C57BL/6 x CBA F1 hybrid mice.

2.1.3 GB mouse lines

GB lines contain both a BLG-loxP and a BLG-Cre transgene. Seven GB lines were produced by systematic breeding of the Lox lines 4, 5, and 9 to the Cre 74 line. Double transgenics were identified by Southern blot probed for both BLG and Cre recombinase. These lines were not maintained.

2.1.4 PLC mouse lines

PLC (PGK-Lox-Cre) lines were derived by microinjecting a circular PGK-Cre construct into pronuclear stage eggs obtained from superovulated C57BL/6 x CBA F1 females mated to lox 4, 5 or 9 males. Microinjection was as per the Lox mouse lines except the construct was injected at 5ng/µl. PCR and Southern blotting were used to identify positive transgenic pups. Transgenic founders were used to establish lines by systematic breeding with C57BL/6 x CBA F1 hybrid mice.

2.2 RECOMBINANT PLASMIDS

2.2.1 BLG-loxP

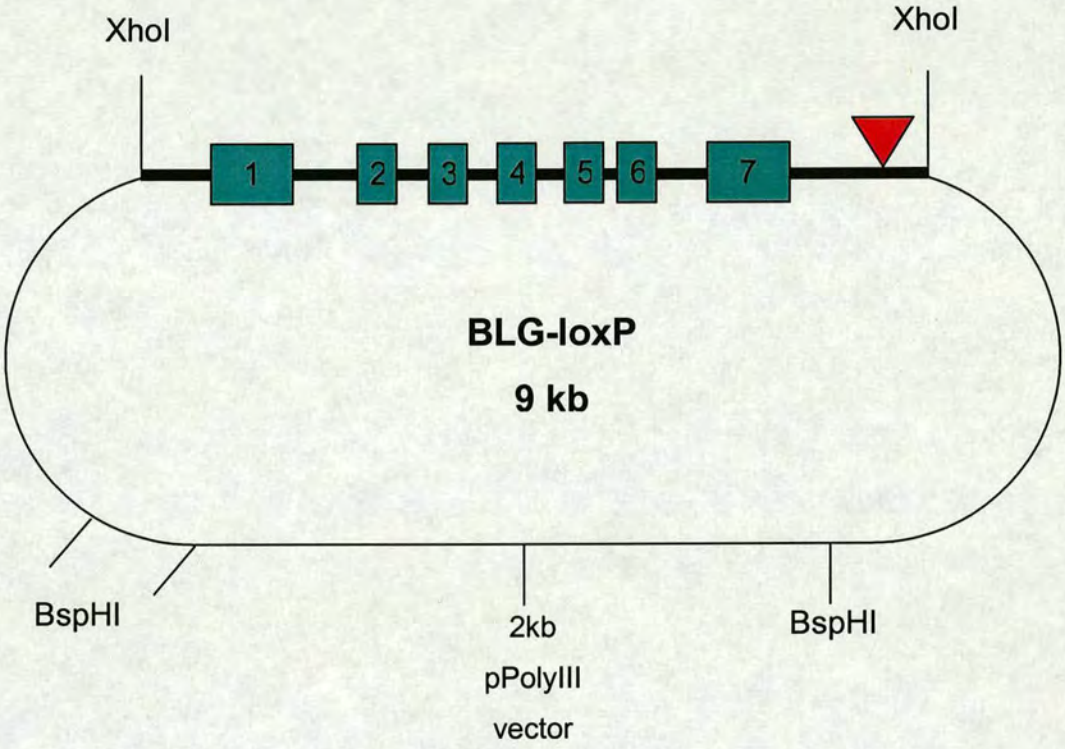


Figure 2.1. BLG-loxP. This is a 9kb plasmid derived from BLGΔDp (Whitelaw *et al* 1992) with the addition of a loxP site at the 3' end. It contains 406bp of 5' flanking region, the 4.7kb BLG structural gene and 1.9kb of 3' regions. Exons are represented by blue numbered boxes, with the whole BLG sequence represented as a thick black line. The loxP site is represented by a red triangle and was inserted at 10320bp (from BLG sequence; corresponds to 6567bp on BLGΔDp), which is 467bp from the 3' end of the BLGΔDp sequence. BspHI sites were used to verify the loxP had integrated. XhoI digestion produces a 7kb BLG fragment that was used to establish BLG-loxP mice by microinjection. The plasmid was supplied by Dr. Doug Strathdee.

2.2.2 PGK-Cre

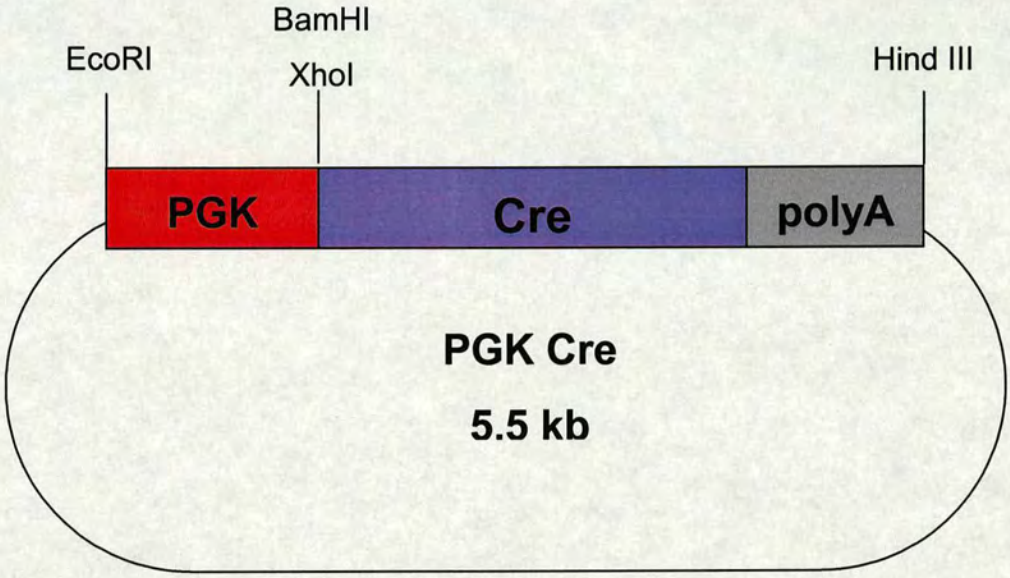


Figure 2.2. PGK-Cre. This is a 5.5kb plasmid containing approximately a 2.7kb fragment containing Cre recombinase and the MT-I(A)n sequence, cloned to a fragment of pSP72 vector and the 800bp PGK promoter. The plasmid was supplied by Dr. Doug Strathdee.

2.2.3 BLG and β-casein expression vectors

The four expression vectors used for production of the riboprobes used for *in situ* hybridisation were supplied by Dr. Maggie McClenaghan (Dobie *et al* 1996).

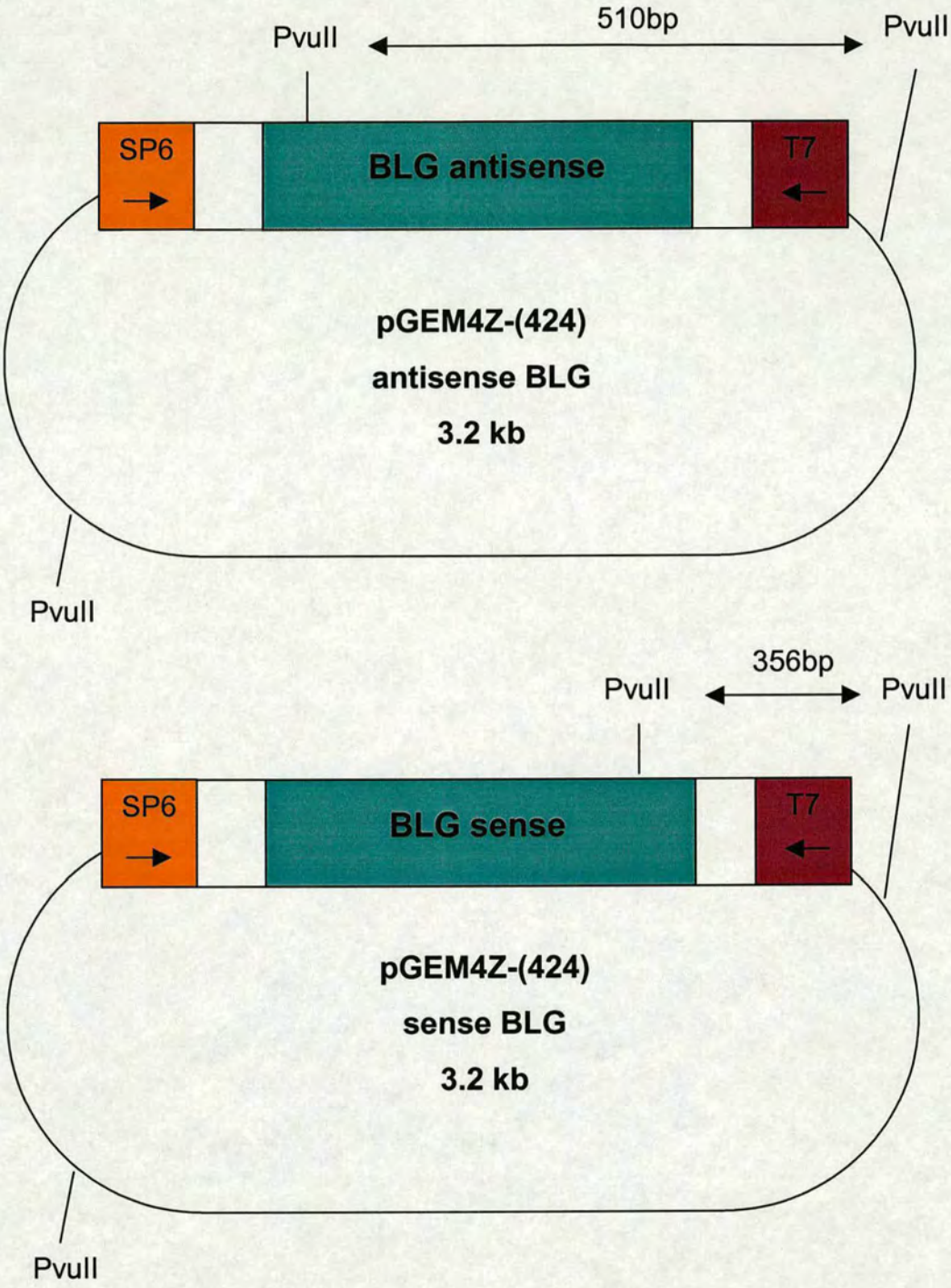


Figure 2.3. BLG expression vectors. pGEM4Z-(424)BLG contains a 424bp fragment of BLG cloned into pGEM-4Z in two directions relative to the T7 promoter (Dobie *et al* 1996). The antisense construct produces a larger PvuII fragment compared to the sense construct. Constructs were linerised with BamHI before use in *in situ* experiments.

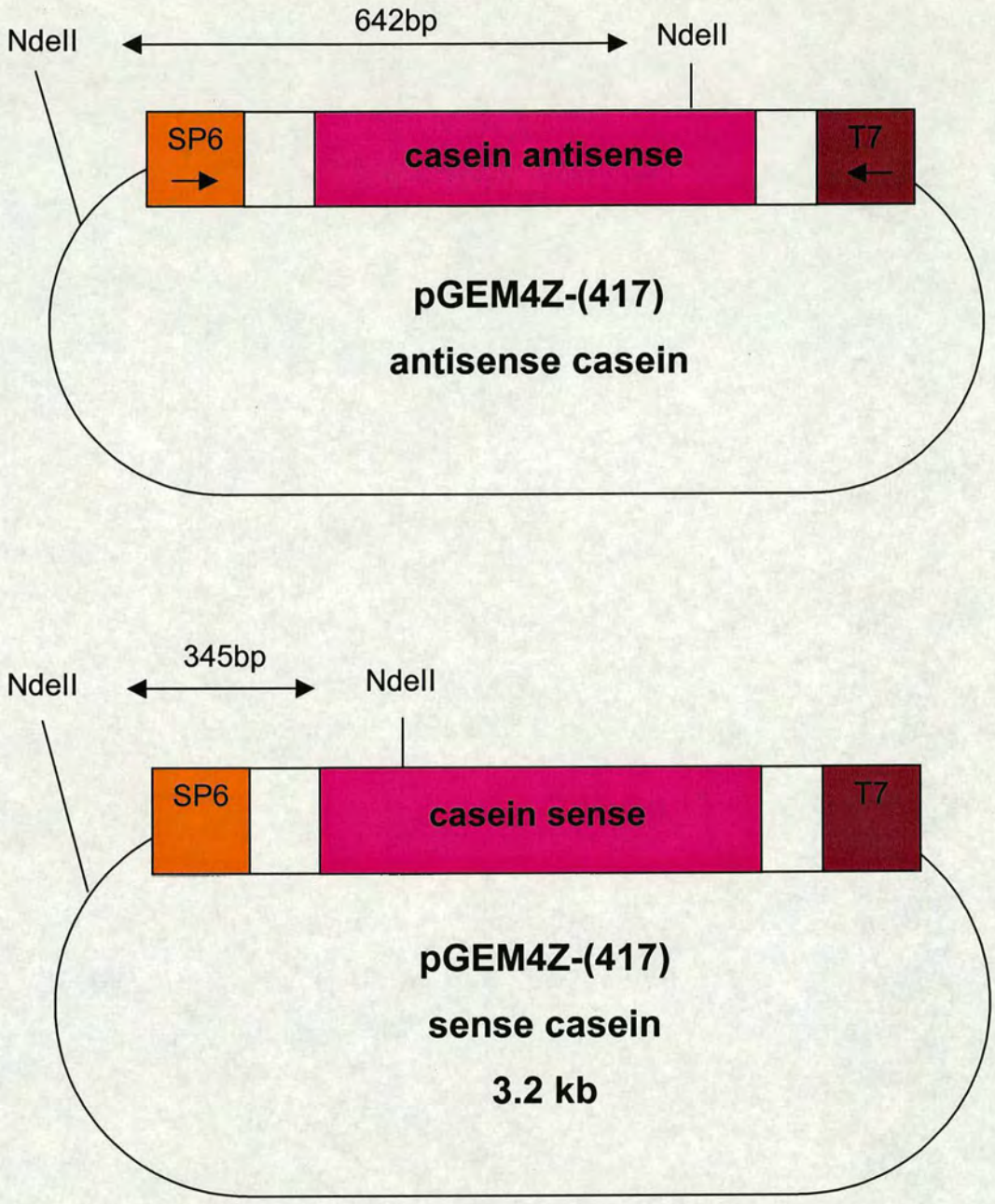


Figure 2.4. Casein expression vectors. pGEM4Z-(417) casein contains a 417bp fragment of β-casein cloned into pGEM-4Z in two directions relative to the T7 promoter (Dobie *et al* 1996). The antisense construct produces a larger NdeII fragment compared to the sense construct. Constructs were linerised with BamHI before use in *in situ* experiments.

2.2.4 Plasmids used for probe production

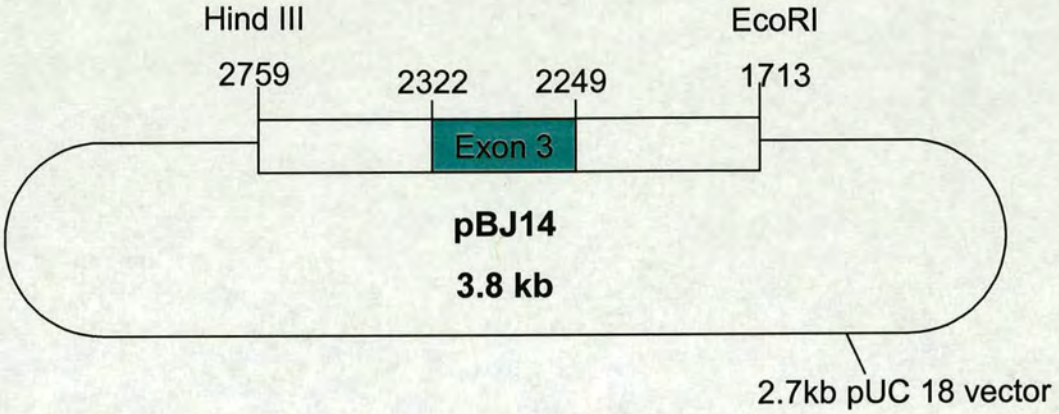


Figure 2.5. pBJ14. This is a 3.8 kb plasmid containing approximately 1.1kb of BLG sequence corresponding to exon III and surrounding introns which was cloned into pUC18. Hind III / EcoRI digestion produces a 1.1kb BLG fragment which was used as a probe for Southern blotting. The figures given for restriction sites are from BLGΔDp sequence. The plasmid was supplied by Dr. Bruce Whitelaw.

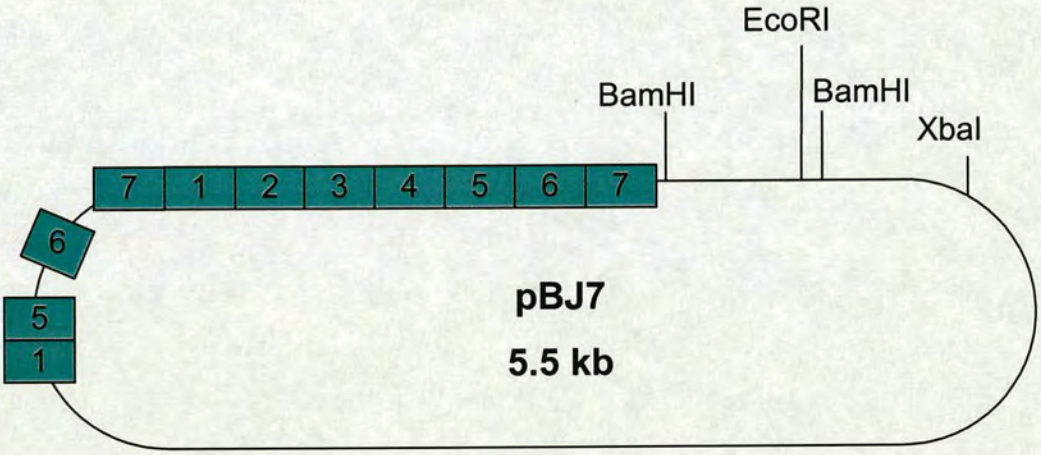


Figure 2.6. pBJ7. This is a 5.5kb plasmid containing the extreme 3' end of the BLG-loxP fragment cloned into ptgpolyIII. Blue boxes denote exons of BLG. EcoRI / XbaI digestion produces a 1kb fragment which was used as a 3' probe for southern blotting.

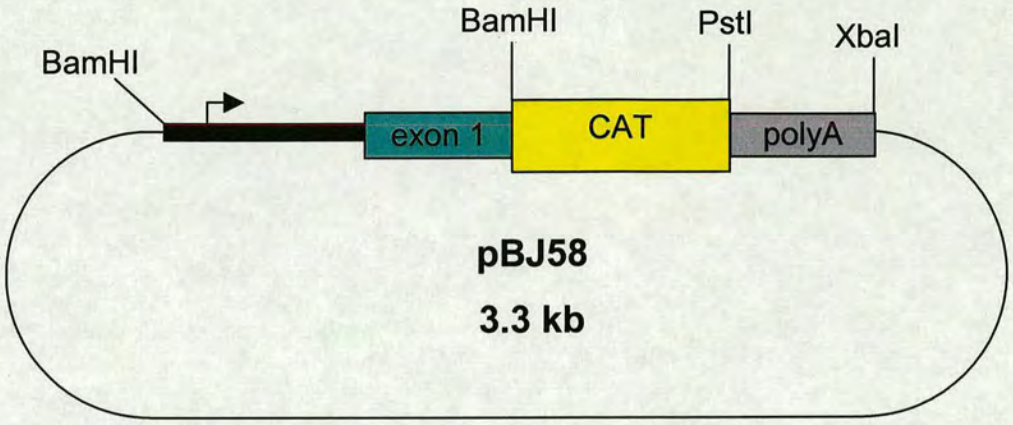


Figure 2.7. pBJ58. This is a 3.3kb plasmid containing approximately 400bp of BLG sequence corresponding to the extreme 5' end of the BLG-loxP fragment which contains the transcription site (denoted by black arrow) and exon I fused to CAT which was cloned into ptgpoly III. BamHI digestion produces a 400bp fragment which was used as a 5' probe for Southern blotting. The plasmid was supplied by Dr. Bruce Whitelaw.

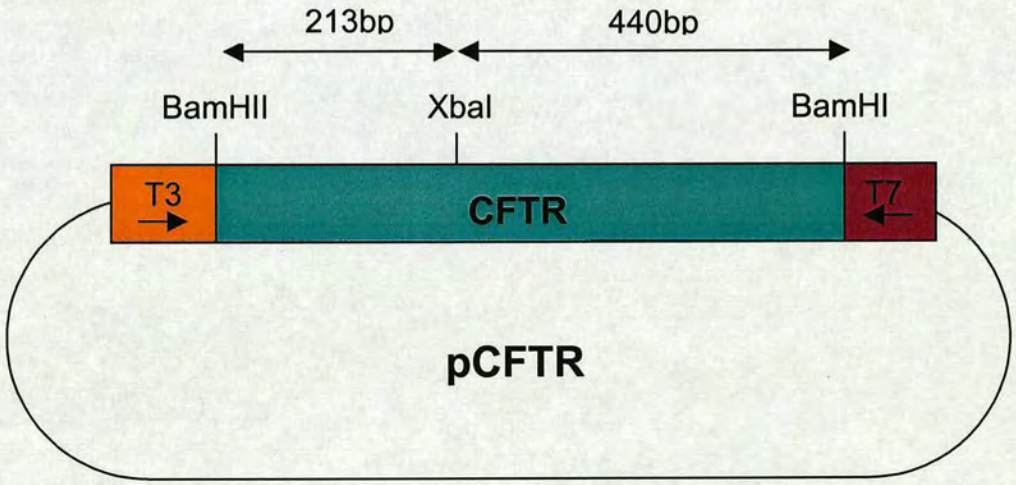


Figure 2.8. pCFTR. This plasmid contains a 653bp fragment of CFTR cloned into Bluescript (M13). BamHI digestion produces a 653bp fragment that was used as a loading control probe for Southern blotting. The plasmid was supplied by Dr.Wei Cui.

A PCR product generated from BLG cDNA was used as a probe for northern blots (supplied by Dr. Romi Pena).

pSX101 contains a 12kb HindIII fragment of rRNA in pAT153 vector (supplied by Dr. Romi Pena). This was used as a loading control for Northern Blots.

2.3 BACTERIAL STRAINS

E.coli DH5α competent cells (Genotype: F^-, ϕ80d*lac*ZΔM15, Δ(*lac*ZYA-*arg*F)U169, *deo*R, *rec*A1, *end*A1, *hsd*R17(r_K^-,m_K^+), *sup*E44, λ-*thi*1, *gyr*A96, *rel*A1) were purchased from GIBCO BRL and stored at –70°C. This strain was used for all plasmid transformation procedures.

2.4 PREPARATION OF MEDIA

2.4.1 Plates containing antibiotics

LB bottom medium (1% w/v tryptone, 0.5% w/v yeast extract, 0.8% w/v agar, 0.1M NaCl) was melted in a microwave and allowed to cool until "hand-hot". 1ml of appropriate antibiotic (e.g. ampicillin; stock = 50mg/ml) was added to 1 litre LB bottom (final concentration of 50μg/ml), swirl mixed and poured immediately into petri dishes with a bunsen flame close by to create a convection current. Once the plates had set, they were placed in a 37°C oven, upside down, with the lids off, for 1 hour to allow excess moisture to evaporate off. Plates were stored at 4°C.

2.4.2 Liquid medium

LB medium (1% w/v tryptone, 0.5% w/v yeast extract, 0.1M NaCl, 10mM $MgCl_2$) was used to grow all *E.coli* strains.

2.5 DNA MANIPULATION AND ANALYSIS

2.5.1 Transformation of DH5α cells with plasmid DNA

DH5α cells were thawed on ice and then gently mixed. 1-3μl (1-10ng DNA) of ligated DNA plasmid was added to 50μl of cells in a microcentrifuge tube and mixed using a pipette tip. The cells were placed on ice for 30 min, heat shocked (37°C for 20 sec) and returned to ice for 2 min. 900μl LB medium (without antibiotic) were added and incubated at 37°C for 1 hour with vigourous shaking. A 1:10 dilution of cells in medium was made and 100μl of each sample was spread onto agar plates (see 2.4.1) using a sterile glass spreader. Plates were incubated at 37°C overnight.

2.5.2 Preparation of plasmid DNA

Plasmid DNA from large-scale bacterial culture (100-500ml) was prepared using Qiagen Plamid Maxiprep DNA purification columns following the manufacturers instructions (Qiagen).

A single colony was lifted using a sterile toothpick and grown at 37°C for 6-8 hours in 5ml LB with 50μl/ml ampicillin. This starter culture was

used to inoculate LB medium for a bulk preparation. For a low copy number plasmid (e.g.pPolyIII derived plasmids) 500ml LB medium with 50μg/ml ampicillin was inoculated. For high copy number plasmids 100ml LB medium with 50μg/ml ampicillin was used. The cultures were incubated overnight at 37°C with vigorous shaking.

Overnight bacterial cultures were decanted into 2 x 250ml bottles and the cells pelleted by centrifugation (5,000 rpm in Sorvall Dry-Spin centrifuge bottles in an F-16 rotor, 10 min at 4°C). After removing the supernatant, the pellet was resuspended by vortexing in 10ml of buffer P1 (50mM Tris-HCl pH8.0, 10mM EDTA, 10μg/ml RNase A). 10ml buffer P2 (200mM NaOH, 1% SDS) was added, mixed gently (the solution should become viscous due to cell lysis), and incubated at room temperature for 5 min. 10ml chilled buffer P3 (3M potassium acetate pH5.5) which precipitates cell debris and chromosomal DNA was added and incubated on ice for 20 min. The bacterial lysate was centrifuged (7,000 rpm as above for 30 min at 4°C). Meanwhile, a Qiagen-tip 500 was equilibrated by applying 10ml Buffer QBT (750mM NaCl, 50mM MOPS pH7.0, 15% isopropanol, 0.15% Triton X-100) and allowing the column to empty by gravity flow.

The supernatant from the centrifuged bacterial lysate was promptly filtered through a prewetted, folded filter and applied to the equilibrated column entering the resin by gravity flow. The column was washed with 2 x 30ml buffer QC (1M NaCl, 50mM MOPS pH7.0, 15% isopropanol) to remove RNA and protein. The DNA was salt-eluted with 15ml buffer QF (1.25M NaCl, 50mM Tris-HCl pH8.5, 15% isopropanol) into a Corex tube containing 10.5ml isopropanol and mixed gently. The sample was centrifuged (8,000 rpm in a Sorvall SS-34 rotor for 30 min at 4°C) and the supernatant removed.

The pelleted DNA was washed with 5ml 70% v/v ethanol and centrifuged (8,000 rpm as above, 10 min at 4°C). After removing the supernatant the pellet was air dried, and resuspended in an appropriate volume of TE.

2.5.3 Preparation of mouse tail genomic DNA

1cm tail-tip biopsies were taken from 3 - 5 week old anaesthetised mice and digested overnight at 37°C in Tail Digestion Buffer (200µg/ml Proteinase K, 0.3M sodium acetate, 10mM Tris-HCl pH7.9, 1mM EDTA, 1% w/v SDS). Samples were vortexed before tissue debris and SDS were pelleted by microcentrifugation (13,000 rpm, 30 min at 4°C). The 700µl supernatant was removed to a tube containing an equal volume of phenol/chloroform and mixed. After microcentrifuging (13,000 rpm, 10 min at room temperature) the extracted aqueous phase was added to an equal volume of isopropanol. The sample was mixed well and left at room temperature for 10 min to allow the DNA precipitate to form. The DNA was pelleted by microcentrifugation (13,000 rpm, 10 min, room temperature), washed with 70% v/v ethanol, air dried and resuspended in 150µl TE buffer. The DNA was stored at –20°C.

2.5.4 Preparation of high molecular weight mouse DNA

Care must be taken when preparing high molecular weight genomic DNA to prevent shearing of the DNA. Shearing was miminised by avoiding vigourous pipetting or vortexing. Large bore pipettes also helped to prevent shearing of the DNA during preparation.

Tissue was frozen in liquid nitrogen immediately after dissection from the animal. Mammary tissue was pulverised and ground to a fine powder under liquid nitrogen in a pestle and mortar before being transferred to a homogeniser due to the nature of the tissue. Liver tissue does not require this step. Two methods were employed during this thesis, because the yield was found to be greater when using method 2 as described below.

Method One

The tissue was added to a 10ml Wheaton homogenizer containing 3ml ice cold RSBE/NP40 (10mM Tris-HCl pH7.4, 10mM NaCl, 2mM EDTA, 0.5% v/v NP40) with 3µl 100mM PMSF added just prior to use. The tissue was liquidised using as few strokes as possible. The homogenate was transferred to a 13ml centrifuge tube and centrifuged (4,500 rpm Jouan CR3000 swing bucket rotor, 5 min at 4°C). The supernatant was removed and the pellet resuspended in 3ml RSBE/NP40 (this wash was repeated twice more). Finally the pellet was resupended in 3ml RSBE (10mM Tris-HCl pH7.4, 10mM NaCl, 2mM EDTA) to which 3ml SNET (1% SDS, 0.6M NaCl, 10mM EDTA, 20mM Tris) was added. Proteinase K was added to a final concentration of 200µg/ml and the viscous solution was incubated at 37°C overnight (or until the solution was no longer cloudy). Two phenol/chloroform extractions and one chloroform extraction were performed (see 2.5.10). The DNA was precipitated with 6ml isopropanol and the "stringy" DNA was spooled out using a heat sealed pipette. The pipette was dipped in 70% v/v ethanol to wash the DNA, followed by air-drying. The DNA was resuspended in 1ml TE and left at 4°C.

Method Two

The tissue was added to a 40ml homogeniser and disrupted in 15ml of 10mM NaCl, 20mM Tris-HCl (pH8.0), 1mM EDTA, 0.5% w/v SDS and 100 μg/ml proteinase K. The solution was transferred to a 50ml falcon tube, mixed and then left overnight at 50°C. The DNA was extracted by gentle inversion with an equal volume of phenol/chloroform for 10 min at room temperature. The DNA was then centrifuged (4000 rpm in an Eppendorf 5810 R Swing bucket rotor, 20 min at 10-12°C). The upper aqueous layer was removed to a clean tube using a wide bore pipette. 1.5ml 3M Sodium Acetate pH5.5 was added and 30ml 100% ethanol was gently layered on top. The DNA was gently mixed for ten minutes by hand. The DNA was spooled out using a heat sealed pasteur pipette, washed in 70% v/v ethanol and air dried for approximately 5 min. The DNA was then resuspended in 2ml of TE and left at 4°C.

2.5.5 Quantifying DNA and RNA

Spectrophotometric measurements were used to determine both the quantity and quality of DNA and RNA prepared. Typically for genomic DNA a 20μl sample was diluted with 980μl distilled water and the absorbance read at 260 nm to give the concentration of nucleic acid in the sample. An OD of 1 corresponds to approximately 50μg/ml of double stranded DNA. An OD of 1 corresponds to approximately to 40μg/ml of single stranded RNA.

The ratio between readings at 260nm and 280nm provides an estimation of the purity of the sample. Pure DNA samples have a

OD_{260}/OD_{280} ratio of 1.8, while pure RNA has a ratio of 2.0. Contamination of a DNA sample with protein produces a lower ratio. The DNA extraction methods used generally gave a ratio of 1.9.

2.5.6 DNA sequencing

The chain termination method was used to sequence part of the BLG-loxP plasmid. A Sequenase Version 2.0 DNA sequencing kit (USB) was used according to the manufacturers instructions and all buffers used as supplied.

This method requires single stranded DNA as its starting material. Therefore, double stranded plasmid DNA was alkaline denatured before use in the sequencing reaction. The DNA was mixed with 0.1 volume of 2mM EDTA and 0.1 volume of 2M NaOH. This was incubated for 30 min at 37°C. 0.1 volume of 3M Sodium Acetate and 2-4 volumes of 100% ethanol were added to the plasmid DNA. This was precipitated at –70°C for 15 min. The DNA was pelleted by microcentrifugation (13,000 rpm, 20 min at room temperature) washed in 70% v/v ethanol, air dried for 5 min and resuspended in 6µl dH_20.

2µl sequenase buffer and 2µl of 20mM primer (sequence: TCCTGTGGGTGTCTGCTTGC) was added to the denatured DNA and the mixture was annealed at 37°C for 30 min. It was then chilled on ice. While the annealing mixture cooled, 4 prelabelled tubes had 2.5µl of termination mixture (either G,A,T, or C) added, and were prewarmed at 37°C. 1µl labelling mix was diluted with 4µl distilled water. 1µl Sequenase

polymerase was diluted with 0.5µl pyrophosphatase and 6.5µl enzyme dilution buffer.

1µl 0.1M DTT, 2µl diluted labelling mix, 0.5µl ^{35}S dATP, 1µl of Manganese buffer and 2µl of diluted sequenase polymerase were added to the 10 µl ice cold annealed DNA mixture. This was incubated at room temperature for 5 min. 3.5µl of this labelling reaction was added to each termination tube (G,A,T,C) and incubated at 37°C for 5 min. The reaction was stopped by adding 4µl of stop solution to each termination tube.

Before loading onto a gel, the samples were heated at 75°C for 2 min, pulse centrifuged, and then 4µl of each reaction was loaded onto a prewarmed denaturing polyacrylamide gel which was run in 1 x TBE for 1-2 hours. After electrophoresis, the gel was removed from the plates by transfer to blotting paper soaked in fixative (5% acetic acid, 20% methanol) and then dried at 80°C for 2 hours. The gel was then exposed to X-ray film (AGFA-Curix Blue).

2.5.7 Digesting DNA with restriction enzymes

All genomic and plasmid DNA was digested using restriction enzymes and buffers according to manufacturers instructions (Roche) except that a 4-10 fold excess of enzyme was generally used. A general example of genomic and plasmid digests is given below. Table 2.1 details the enzymes and buffers used in this study.

Example 1	Incubated at 37°C overnight
10µg	genomic DNA (in 30µl distilled water)
2µl	enzyme 40U/µl (Roche)
5µl	10 x buffer (Roche)
1µl	Spermidine 60mM
1µl	BSA 10 mg/ml (if required for the enzyme)
to 50µl	distilled water

Example 2	Incubated at 37°C for 1-2 hours
1µg	plasmid DNA (in 5µl distilled water)
1µl	enzyme 10U/µl (Roche)
2µl	10 x buffer (Roche)
to 20µl	distilled water

Enzyme	*Buffer*	*BSA recommended*
BspHI	B	No
Eco RI	H	No
Eco RV	B	No
Hind III	B	No
Kpn I	L	Yes
Nhe I	M	No
Not I	H	No
Ssp I	H	No
Xba I	H	No
Xho I	H	No

Table 2.1 Restriction enzymes and conditions for digestion. All enzymes and buffers used were purchased from Roche. Buffer B=10mM Tris-HCl, 5mM $MgCl_2$, 100mM NaCl, 1mM 2-Mercaptoethanol, pH 8.0. Buffer H=50mM Tris-HCl, 10 mM $MgCl_2$, 100mM NaCl, 1mM DTE, pH 7.5. Buffer L=10mM Tris-HCl, 10mM $MgCl_2$, 1mM DTE, pH 7.5. Buffer M=10mM Tris-HCl, 10mM $MgCl_2$, 50mM NaCl, 1mM DTE, pH 7.5

2.5.8 Running agarose gels

Horizontal, buffer-immersed agarose gels containing EtBr (0.25µg/ml) were run in 1 x TAE buffer. Two sizes of gels were run: 8 x 10 cm mini gels were used for probe preparation, while 20 x 20 cm gels were used for Southern blotting. Agarose concentration varied between 0.4 –2 % w/v depending of the size of fragments to be resolved. Before loading, 0.1 volume of buffer II (Sambrook *et al* 1989) was added to each sample. Gels were run at varying voltages: Southern blot gels were usually run overnight at 30V, BLG PCR gels were run at 100V for 30 min, while enzyme restriction digests were run at 100V for 1 hour. DNA was visualised using an UV transilluminator and images were captured using Multi-Analyst software (BioRad).

2.5.9 Purifying DNA fragments from agarose gels

DNA fragments were electrophoresed through an agarose gel, the fragment was excised from the gel with a clean sharp scalpal (visualised using long wavelength ultraviolet light) and transferred to a microcentrifuge tube. The gel slice was then processed using a QIAquick column (Qiagen) according to the manufacturers instructions. The weight of the agarose was determined and 3 volumes of buffer QG to 1 volume of gel was added (e.g. 300µl buffer QG to 100mg gel). This was incubated at 50°C for 10 min to dissolve the gel slice. 1 gel volume of isopropranol was added, mixed and the sample applied to a QiaQuick column held within a 2ml collection tube. After 1 min microcentrifugation the flow through was discarded and the column recentrifuged (1min) with an empty collection tube. The column was washed with 0.75ml buffer PE by 1 min microcentrifugation, flow

through was discarded and the column recentrifuged (1 min). The column was then placed in a clean 1.5ml microcentrifuge tube, 30µl elution buffer (10mM Tris-HCl pH 8.5) was added, allowed to stand for 1 min, and then microcentrifuged for 1 min to elute the DNA.

2.5.10 Phenol/Chloroform extraction of DNA

Phenol/chloroform extraction is used to deproteinise the aqueous phase containing the DNA. Mixing phenol with the DNA sample causes the dissociation of proteins from DNA. After centrifugation two phases are present: a lower organic phenol phase carrying the protein (much of which is found in the white interphase) and the aqueous phase which contains the DNA. Chloroform improves the efficiency of extraction due to its ability to denature proteins. The high density of chloroform also enhances the separation of phases facilitating the removal of the aqueous phase with little cross-contamination from organic material.

An equal volume of phenol (equilibrated to pH7.6 by repeated extraction with Tris buffer; Fisher Scientific UK) was added to each sample in microcentrifuge tubes, mixed thoroughly to form an emulsion, and microcentrifuged (13,000rpm, 10 min at room temperature) to separate the organic and aqueous phases. The top aqueous phase was transferred to a clean tube taking care not to disturb the interphase. The process was repeated using a 1:1 mix of phenol/chloroform followed by chloroform (this removes any residual phenol which would inhibit future enzymatic reactions).

DNA was then precipitated with 0.1 volume of 3M Sodium Acetate pH5.5 and 2 volumes of 100% ethanol, mixed and then chilled to –70°C for 30 min. The DNA was pelleted by microcentrifugation (13,000 rpm, 20 min at room temperature). The supernatant was removed and the DNA pellet washed twice (centrifuged for 2 min between washes) with 70% v/v ethanol to remove salts. The pellet was then air dried and resuspended in TE. Care was taken to dry the pellet sufficiently to ensure that all ethanol was removed since ethanol can inhibit certain enzymatic reactions. Care was also taken to ensure that the pellet was not too dry, as the DNA can be very difficult to solubilise, especially with genomic DNA. Generally, adding TE buffer when the edges of the pellet become "glass like" avoided the problem. Samples were stored at –20°C.

2.5.11 Southern Blotting

2.5.11.1 Transfer onto Zetaprobe GT membrane

Southern blotting allows the transfer of DNA from an agarose gel onto a membrane, which can then be analysed for the presence of specific sequences (Southern 1975). This aim can be achieved by the flow of liquid through the gel and membrane. The DNA is carried with the liquid but cannot pass through the membrane and this is followed by fixation of the DNA to the membrane.

If the fragments are >4kb, the gel was soaked in denaturing solution (0.5M NaOH, 1.5M NaCl) for 2 x 20 min, neutralisation buffer (0.5M Tris-HCl pH8.0, 3M NaCl) for 2 x 20 min to facilitate transfer of all DNA regardless of

original size. Air bubbles between layers of the stack were avoided as they inhibit transfer. The gel was then placed on a sheet of 3MM Whatman paper (saturated in 0.4M NaOH solution) with the ends of the paper resting in a reservoir of 0.4M NaOH. Saran wrap was placed over the whole stack and a window was made, allowing only the gel to emerge (this avoids 'short circuiting' ensuring capillary action occurs only through the gel). Zetaprobe GT membrane (Bio-Rad) was cut to the size of the gel, prewetted in distilled water, and then placed carefully on top of the gel and flooded with 0.4M NaOH. Two precut, prewetted (distilled water) 3MM sheets were placed on top, followed by two dry 3MM sheets. A stack of paper towels was then placed on top, followed by a glass plate to keep all the layers together and finally a small weight. Transfer was allowed to occur overnight.

After transfer the stack was disassembled down to the membrane, the wells were marked with a soft pencil and the membrane was washed briefly in 2 x SSC. The DNA was fixed to the membrane by oven drying at 70°C for at least 1 hour.

2.5.11.2 Transfer onto Hybond – N membrane

The process for transfer to Hybond-N membrane (Amersham Life Sciences) was identical to that described for Zetaprobe GT membranes except that 20X SSC was used as the transfer buffer.

2.5.12 Preparation of radiolabelled DNA probes and hybridisation to Southern Blot membranes

2.5.12.1 High Prime DNA labelling method

The High Prime (Roche) labelling method is based on "random primed" DNA labelling originally developed by Feinberg and Vogelstein (1984). It is based on the hybridisation of oligonucleotides of all possible sequences to the denatured DNA. The complementary DNA strand is synthesised by Klenow polymerase and modified deoxyribonucleoside-triphosphates present in the reaction are incorporated into the newly synthesised cDNA strand.

High Prime labelling mix contains: random primers along with 1U/µl Klenow polymerase, 0.125mM dATP, 0.125mM dGTP, 0.125mM dTTP and 5x reaction buffer in 50% v/v glycerol. 25-50ng of probe DNA was made up to 11µl with distilled water and boiled for 10 min to denature the DNA. The mixture was spun briefly and then placed on ice for 2 min. 4µl High Prime labelling mix was added followed by 5µl of [α^{32}P]dCTP [~110TBq/mmol]. The sample was mixed gently, spun briefly and then incubated at 37°C for 10 – 20 min.

To help with size estimation Southern blots were often probed with a labelled molecular weight ladder e.g. lambda cut with HindIII. To prepare this 25ng of ladder DNA was made upto 11µl with distilled water and boiled for 10 min. The mixture was spun briefly and then placed on ice for 2 min. 1µl of [α^{32}P] dCTP [~110TBq/mmol] and 9µl of 'cold' dCTP were added,

followed by 4µl High Prime labelling mix. This ensured the signal from the ladder did not swamp that from the genomic samples. The sample was mixed gently, spun briefly and then incubated at 37°C for 10-20 min.

2.5.12.2 Purifying labelled DNA probe

After the labelling of the probe unincorporated nucleotides were removed from the labelled probe using a Sephadex G50-Nick Column (Pharmacia Biotech). Removal of unincorporated nucleotides helped reduce background signal.

The storage buffer was removed and the column was equilibrated by adding 2.5ml TE buffer and allowing it to enter the resin by gravity flow. The 20µl sample was added directly to the column and washed through with 400µl TE buffer. This eluate was discarded. A further 400µl TE buffer was used to elute the labelled probe. The sample was then quantified.

2.5.12.3 Quantifying ^{32}P-labelled DNA probes

2µl of the 400µl purified probe mixture was spotted onto 2.5cm glass microfibre filter paper (Whatman) and placed in a scintillation vial (without scintillation fluid). A Wallac 1410 β-counter (Pharmacia Biotech) was used to analyse β-emission levels. Probes with a total activity of at least 10^8 cpm were used.

2.5.12.4 Hybridisation, washes and detection of probes

Prehybridisation of membranes was performed in revolving glass cylinders (Techne) using 20ml prewarmed hybridisation solution (0.25M disodium hydrogen orthophosphate [Na_2HPO_4] pH7.2, 7% w/v SDS, 1mM EDTA) in a Techne Hybridiser HB-1D oven set at 65°C for at least 10 min.

Prior to hybridisation, the probe was denatured with 30µl 5M NaOH. Denatured probe was added to the prehybridised membranes and hybridised with rotation overnight at 65°C.

Membranes were washed to remove unhybridised probe and background signal: 2 x 20 min in 50 ml low stringency wash (20mM disodium hydrogen orthophosphate [Na_2HPO_4] pH7.2, 5% w/v SDS), followed by 2 x 20 min in 50 ml high stringency wash (20mM disodium hydrogen orthophosphate [Na_2HPO_4] pH7.2, 1% w/v SDS). The membrane was wrapped in Saran wrap to avoid it drying out between probings.

Membranes were either exposed to Phosphor Screens (Bio-Rad) at room temperature (1 – 72 hr) or to X-ray film (AGFA-Curix Blue) between intensifying screens in autoradiography cassettes at –70°C (6hrs – 2 weeks). Phosphor Screens were processed using a Molecular Imager FX (Bio-Rad) and X-ray film was developed using an X-ograph Compact X2.

If necessary, the membranes were stripped by placing them in boiling 0.1% w/v SDS and allowing them to cool to room temperature before rewrapping in plastic wrap. The membrane was then ready for reprobing.

2.5.13 Polymerase chain reaction analysis of tail biopsy DNA

A PCR assay was used for rapid screening for the presence of the BLG-LoxP transgene during breeding programmes. The bench and pipettes were wiped with 100% ethanol and a separate bottle of sterile distilled water was used for PCR reactions to reduce the risk of DNA contamination leading to false positives.

1µl of centrifuged supernatant from digested tail biopsy was added to 25µl of PCR reaction mix. If a Hybaid Omnigene PCR or Techne PHC-2 thermocycler was used the reaction was overlayed with a drop of mineral oil, if using a Perkin Elmer GeneAmp 9700 thermocycler with heated lid, no oil was needed. The PCR reaction mix comprises the following:

2.50µl	10x buffer (Roche)
0.75µl	50mM $MgCl_2$ (Roche)
0.25µl	20mM BLG primers §
0.25µl	20mM HPRT primers ‡
2.50µl	DMSO
4.00µl	1.25mM dNTPs
14.60µl	distilled water
0.15µl	*Taq* DNA polymerase (Roche)

§BLG primers:	5'-GCT TCT GGG GTC TAC CAG GAA C-3'
	5'-TCG TGC TTC TGA GCT CTG CAG-3'
‡HPRT primers:	5'-GAG TTC CGG AAC TGC CTT TGG TG-3'
	5'-CTG TGC CAC CGG GCG CAT GG-3'

In both cases equimolar amounts of both primers are added together to make a 20µM working stock

Samples were placed in the PCR thermocycler:

Hybaid/Techne		**Perkin Elmer**	
30 x	92°C for 30 sec	30 x	94°C for 1.5min
	65°C for 5 min		65°C for 2 min
		1 x	72°C for 7 min

Once the reaction was complete 10µl of PCR loading dye (15% w/v Ficoll 400, 0.05% w/v SDS, 20mM EDTA pH8.0, 0.125% w/v Orange G) was added to each sample, mixed and 30µl electrophoresed in 2% w/v agarose gels. BLG primers amplify a 248bp segment of the 5' end of the BLG gene, while the HPRT primers serve as an internal control by amplifying a 332 bp segment of the HPRT gene. Control positive, negative and blank samples were set up with every set of diagnostic PCR reactions.

2.5.14 Long Range PCR

A long range PCR sytem was used to amplify the transgene array from PLC A mouse DNA. A mix of thermostable *Taq* polymerase and a proofreading *Pwo* polymerase provide the means to amplify large genomic fragments. Expand Long Template PCR System (Roche) was used according to the manufacturers instructions. Pipettes were washed with 70% ethanol before used and the PCR reaction mixes were made up in a fumehood using filtered tips to avoid contamination. Distilled water and negative DNA was always used as negative controls.

The following was mixed together in an microfuge tube labelled A:

10mM dATP	2.5µl
10mM dCTP	2.5µl
10mM dGTP	2.5µl
10mM dTTP	2.5µl
5′ junction primer 10µM	1.5µl
nested3′ primer 10µM	1.5µl
DMSO	1µl
Distilled water	10µl

5′ junction primer:	5′-GAA TTC GAG CTC GGTACC CTC CCT TCA-3′
nested 3′ primer:	5′-TAG AGA GAA GCC TGC GCA CCG AAC-3′

The following was mixed together in another microfuge tube labelled B:

Enzyme buffer 3 (Roche)	5µl
PCR enzyme (Roche)	0.75µl
Distilled water	15.25µl

24µl of mix A was added to a 0.2ml thin walled PCR tube, mixed with 1µl of the sample, and 25µl of mix B was added. The whole reaction was mixed by pipette action. 30µl oil was added over the reaction and lids sealed over the tubes. The tubes were then placed in a Perkin Elmer GenAmp 9700 and cycled as below:

1x	2 min at 92°C
10x	10 sec at 92°C
	30 sec at 65°C
	8 min at 68°C
30x	10 sec at 92°C
	30 sec at 65°C
	8 min + additional 20 sec per cycle at 68°C
1x	7 min at 68°C

Once the reaction was complete, 25µl was electrophoresed on 0.7% agarose gels and then photographed.

2.6 MILKING OF MICE

All milk samples were collected at midlactation (day 11). Litters were standardised to 5 pups per mother at birth and were separated from their mother 3 hours before milking. Mothers were injected with 0.3 International Units of oxytocin (Sigma) in 0.1ml of distilled water into the intraperitonal space. After 10 min an injection of anaesthetic was given (Hypnorm/Hypnovel; 10µl/g). After 10 min gentle massage was applied to the mammary glands and milk collected in 50µl capillary tubes. Milk was dispensed into a microcentrifuge tube (typically 250µl per mouse) and stored at –20°C.

2.7 MILK PROTEIN ANALYSIS

2.7.1 Preparation of milk proteins

Defatted milk was prepared by adding 120µl distilled water to 30µl thawed mouse milk, briefly vortexing, followed by a pulse spin in a microcentrifuge. Care was taken to disrupt any protein pellet that formed without disturbing any of the fat around the side of the tube. The defatted milk (supernatant) was transferred to a clean tube. A 50µl sample of the defatted milk was further diluted in 200µl reducing buffer (0.5M Tris-HCl pH6.8, 10% v/v glycerol, 10% w/v SDS, 5% v/v β-mercaptoethanol, 0.1% w/v bromophenol blue). The samples were boiled for 4 minutes to denature the proteins. Any lost volume due to evaporation was replaced and a last 0.1 dilution (e.g. 10µl sample and 90µl reducing buffer) gave a final dilution of 1/250. Samples were stored at –20°C.

2.7.2 Preparation of SDS-polyacrylamide gels

Quantitative milk protein analysis was performed using SDS polyacrylamide gel electrophoresis (SDS-PAGE). Gels were either hand poured using Bio-Rad equipment or pre-poured 18%/4% Tris-HCl Criterion gels (BioRad).

If gels were hand poured the following procedure was followed. 20 x 20 cm, 20 x 18 cm glass plates and 1mm gel spacers were thoroughly cleaned using 70% v/v ethanol. They were then clamped into a vertical position to act as a gel mould. A 17.5% resolving gel was prepared by mixing 4.75ml distilled water, 7.5ml 1.5M Tris-HCl pH8.8, 300µl 10% w/v SDS, 17.5ml

Amerseco acrylamide solution (30% w/v acrylamide – 1.034% w/v bis-acrylamide, a ratio of 37.5:1). The gel was polymerised by adding 150µl of fresh 10% w/v APS and 15µl TEMED, mixed and poured immediately into the 1mm space between the glass plates. A one cm layer of distilled water was layered on top to aid polymerisation and help produce a smooth edge to the gel. The resolving gel was usually allowed to polymerise overnight. Once the resolving gel had set the water layer was removed and a 4% stacking gel was prepared. 4ml of distilled water, 1.66ml 0.5M Tris-HCl pH6.8, 67µl of 10% w/v SDS, 866µl Ameresco acrylamide solution (see above) were mixed, then 34µl of 10% w/v APS and 4µl TEMED were added. The gel was poured on top of the resolving gel and a comb was placed in the stacking gel to produce wells before polymerisation occurred. The stacking gel was allowed to polymerise for 1 hour. The comb was removed and the wells washed thoroughly with gel running buffer (25mM Tris, 0.2M glycine, 0.4mM SDS) before samples were loaded.

2.7.3 Electrophoresis of milk proteins

50µl of diluted defatted mouse milk in reducing buffer was loaded into polyacrylamide gels using a Hamilton syringe. After loading the hand poured gels were placed into Bio-Rad Protean II running apparatus with ~ 4l of gel running buffer in the lower reservoir and ~300ml in the upper reservoir. Samples were electrophoresed at 30mA (constant current) while in the stacking gel (~1 hour) and at 60mA while in the resolving gel (~5 hours). Criterion gels were placed in a Criterion gel tank (BioRad) with ~ 1l of gel running buffer in the lower reservoir and ~50ml in the upper reservoir and run at 100V for approximately 3 hours.

Milk protein gels were washed with gentle agitation overnight in fix/stain solution (50% v/v methanol, 7.5% v/v acetic acid, 2% w/v TCA, 0.04% w/v Coomassie blue). Background stain was removed by washing with gentle agitation for 2 x 30 min in destain solution (23% v/v methanol, 7% v/v acetic acid) or until background staining was removed. Gels were placed on glass plates, wrapped in Saran wrap and stored at 4°C. Gels were analysed using Multi-Analyst software (BioRad). Milk BLG protein concentrations were standardised against the albumin band to control for loading differences.

2.7.4 Western Blotting of protein gels

After electrophoresis on SDS-PAGE gels, proteins were electrophoretically transferred to nitrocellulose membranes (Schleicher and Schuell) using the semi-dry method. A stack was constructed: three sheets of Whatman 3MM paper (soaked in transfer buffer: 48mM Tris, 39mM Glycine, 20% methanol, 1.3mM SDS, pH 9.2) onto the lower graphite anode plate, one sheet of nitrocellulose membrane (soaked in distilled water), the gel (soaked in transfer buffer), three sheets of Whatman 3MM paper (soaked in transfer buffer), followed by the upper graphite cathode plate. Transfer was performed at 20V for 1 hour in a BioRad Trans-Blot SD Semi-Dry Transfer Cell.

After transfer the nitrocellulose membrane was washed in TBS-T (100mM Tris pH7.5, 150mM NaCl, 0.1% Tween 20) with 2 % BSA for 2 hours at room temperature, incubated with a rabbit antibody against BLG (a gift from Dr. Maggie McClenagahan) [1/200 dilution in TBS-T 1% BSA, total volume 15 ml] for 2 hours at room temperature, washed in TBS-T for 3 x 10

min, incubated with Horseradish peroxidase-conjugated donkey anti-rabbit immunoglobulin (Diagnostics Scotland) [1/200 dilution in TBS-T 1% BSA, total volume 15ml] for 1 hour at room temperature, followed by 3 x 10 min washes in TBS-T. The secondary antibody was detected using ECL (Amersham) according to the manufacturers instructions. In darkroom conditions 8 ml each of reagent 1 and 2 are mixed, added to membrane for one min and then drained off. The membrane is wrapped in Saran wrap and exposed to X-ray film for the required time.

2.8 TISSUE *IN SITU* HYBRIDISATION

2.8.1 BLG and β-casein expression vectors

Riboprobes of between 200bp and 1kb are recommended for tissue in situ experiments using ^{35}S labelled probe and paraffin embedded tissue. Probes less than 200bp give a reduced hybridisation signal and probes above 1kb penetrate less effectively into fixed tissue (Angerer and Angerer 1992).

A 424bp *Pst* I BLG cDNA fragment consisting of all of exons 3-6 and 6bp of exon 2 and 81bp of exon 7 had been cloned in both orientations relative to the T7 promoter in pGEM-4Z (Dobie *et al* 1996). Cloning the inserts in both orientations allows synthesis of anti-sense (probe) and sense (control for non-specific background) riboprobes from the same promoter. An off centre internal *Pvu* II was used to check the orientations.

A 417bp *Pst* I β-casein cDNA fragment consisting of 5bp of exon 6 and 412bp of exon 7 had been cloned in both orientations relative to the T7 promoter in pGEM-4Z (Dobie *et al* 1996). An off centre internal Nde I site was used to check the orientations.

Plasmid DNA was prepared for each of the four vectors (2.5.2). The expression vectors were linerized using a *Bam* HI site (downstream of the inserts and the T7 promoter) so that transcripts were free from vector sequences (to reduce isotope usage and decrease background hybridisation), phenol/chloroform extracted, ethanol precipitated and stored in DEPC-treated distilled water at –20°C.

2.8.2 Preparation of tissue

Paraformaldehyde is a cross-linking fixative used to retain mRNA in sections during hybridisation and wash procedures.

Tissue was fixed in 4% w/v PFA in PBS pH7.2 overnight at 4°C with gentle agitation. The tissue was washed in PBS for 30 min at 4°C, 30% v/v ethanol for 15 min at room temperature, 70% v/v ethanol for 30 min at room temperature, followed by a final 70% v/v ethanol wash for 30 min at room temperature. The tissue could be stored (4°C) at this stage for upto several months with no noticeable deterioration in mRNA. The tissue was then transferred to a histocasette, labelled with pencil and further processed using a Shannon HyperCentre as follows:

70% v/v ethanol	room temp	vacuum	30 min
70% v/v ethanol	room temp	vacuum	30 min
640P*	35°C	vacuum	1.5 hr
640P	35°C	vacuum	1.5 hr
740P*	35°C	vacuum	1.5 hr
740P	35°C	vacuum	1.5 hr
740P/CNP 30	35°C	vacuum	1.5 hr
CNP 30	35°C	vacuum	1.5 hr
CNP 30	40°C	vacuum	1.5 hr
CNP 30	45°C	vacuum	1.5 hr
Paraffin wax	53°C	vacuum	1.5 hr
Paraffin wax	60°C	vacuum	1.5 hr

* 640P is 64% over proof and is equivalent to approx 96% v/v alcohol

*740P is 74% over proof and is equivalent to approx 99% v/v alcohol

The tissue was finally embedded in a mould containing fresh molten paraffin wax and allowed to set. The tissue blocks were stored at 4°C (the tissue can be stored this way for several years).

2.8.3 Coating slides with TESPA

Tissue sections need to be fixed to the slides, otherwise they float off during subsequent processing. This is achieved by coating the slides with TESPA which acts as a tissue adhesive. The hydrophobic silane moiety of TESPA binds to the glass slide, while the amino group (activated by aldehyde) is crosslinked to the tissue section.

Slides were placed horizontally in glass racks and oven baked overnight at 180°C. Batches of slides were washed for 10 sec in (10% v/v HCl, 75% v/v ethanol), DEPC treated distilled water for 10 sec, 100% acetone for 10 sec and air dried. The slides could be stored in a dust free environment for months at this stage. The slides were coated with TESPA by washing in 2% v/v TESPA in acetone for 10 sec, followed by 2 washes in 100% acetone for 10 sec. The slides were air dried and stored in a slide box with silica gel for up to a month before use.

2.8.4 Tissue sectioning

The paraffin blocks of embedded tissue were kept chilled to 4°C. Any instruments that would come into contact with the tissue including the rotary microtome, steel blades, manipulation utensiles and empty water bath were wiped down with 100% ethanol. The water bath was filled with ~2l of DEPC-treated distilled water and heated to 42-45°C. After trimming the blocks, 5μm sections were cut to generate a ribbon of sections which were placed on a clean untreated glass slide. 30% v/v ethanol was added underneath the ribbon to avoid creases in the sections. The slide was gently dipped into the water bath and the ribbon allowed to float off. The wax was allowed to melt slightly (~1 min) so that the tissue sections expanded to their original dimensions. TESPA coated slides were dipped vertically into the water to "catch" the section. Tissue sections were arranged onto slides in sets of 4 abutting sections. The slides were dried overnight in an oven at 60°C. Slides were stored in a slide box containing silica gel at 4°C. The slides could be stored at this stage for a number of months.

2.8.5 Prehybridisation

The aim of this step is to partially remove proteins in order to increase target RNA accessibility to probes and to decrease background. Proteinase K is used to remove proteins, xylene removes residual paraffin (which traps probe), and acetic anhydride helps decrease positive charges on sections thereby reducing background signal.

Slides were placed in a glass rack and treated with 200ml of the following solutions using gentle agitation: 100% xylene for 2 x 5 min (in fume hood), through a rehydrating ethanol gradient of 2 x 100%, 90%, 70%, 50% and 30% v/v ethanol solutions for 2 min each, PBS for 2 min, 4% w/v PFA in PBS pH7.2 for 10 min, PBS for 2 x 2 min, proteinase K solution for 7.5 min (50mM Tris, 5mM EDTA, 5µg/ml proteinase K from *Tritirachium album* [Sigma]), PBS for 1 min, 4% w/v PFA in PBS pH7.2 for 2 min, DEPC-treated distilled water for 10 sec, 0.1M TEA pH8.0 for 30 sec, 0.1M TEA with 625µl acetic anhydride for 2 x 5 min (acetic anhydride added immediately before use; to inactivate proteinase K), PBS for 2 min, 0.85% w/v NaCl for 2 min, through a dehydrating ethanol gradient of 30%, 50%, 70%, 90% v/v ethanol solutions for 2 min each followed by 100% ethanol for 3 x 5 min. The slides were air dried and then stored in a slide box (slides were generally hybridised the same week).

2.8.6 Generation of ^{35}S labelled riboprobes

^{35}S rUTP labelled single-stranded antisense RNA probes and control sense RNA probes were generated by *in vitro* transcription from the T7 promoter using a Riboprobe *in vitro* Transcription kit (Promega).

The following mixture was prepared in a microcentrifuge tube at room temperature in listed order:

6µl	5x Buffer (Promega)
1µl	10mM rATP
1µl	10mM rCTP
1µl	10mM rGTP
1µl	1M DTT
12µl	^{35}S rUTP
5µl	0.5 – 1 mg/ 5ml DNA template
1.2µl	RNase block (40U/µl; Promega)
0.8µl	T7 polymerase (19U/µl; Promega)

The samples were vortexed briefly and pulse-centrifuged before incubating at 37°C for 25 min. Another 0.8µl T7 polymerase was added and then incubated at 37°C for a further 25 min. 2µl tRNA (stock = 10mg/ml, Sigma) and 4µl DNase (1U/ul) were added and incubated at 37°C for 10 min. 2µl of 100mM EDTA and 162µl of TE buffer containing 50mM DTT were added giving a final volume of 200µl. The riboprobes were purified by two phenol extractions and one phenol/chloroform extraction and precipitated at –20°C overnight with ethanol (see 2.5.10).

The precipitated riboprobes were pelleted by microcentrifugation (13,000 rpm, 30 min at room temperature). Meanwhile, an 80% v/v ethanol, 50mM DTT wash solution was prepared (100µl 1M DTT, 1.9ml TE buffer and 8ml 100% ethanol). RNA pellets were washed twice with 500ml of the wash solution and then once in 100% ethanol, air dried and resuspended in 25µl of 50mM DTT, TE buffer.

2.8.7 Quantification of ^{35}S labelled riboprobes

1µl of the probe was added to 19µl distilled water and 1µl of the dilution was spotted onto each of two 2.5cm glass microfibre filters (Whatman). One filter was washed three times with 5% w/v TCA and once with chilled 100% ethanol. The filters were placed in scintillation vials and a drop of scintillation fluid was added. The filters were then analysed using a Wallac 1410 β-counter (Pharmacia Biotech). The percentage incorporation was measured from the two filters using the counts per minute values. CPM were converted to disintegration per minute by assuming 50% counting efficiency; therefore CPM values were multiplied by two to give DPM values. An incorporation of 1 x 10^5 dpm per 1µl probe was the target level of activity. To achieve this the probe was diluted with hybridisation solution. Given an average incorporation value four probes would require about 5ml hybridisation solution. Hybridisation buffer was made at room temperature and comprises the following:

50% formamide:	to reduce the hybridisation temperature
10% dextran sulphate:	increase hybridisation rate
1 x Denhardt's solution:	reduces non-specific binding
20mM Tris-HCl pH8.0	
0.3M NaCl:	reduces electrostatic binding of probe to tissue
5mM EDTA	
10mM sodium phosphate	
0.5mg/ml tRNA:	carrier RNA
50mM DTT:	prevents probe from oxidising and inhibits Rnase

Probes were stored at –20°C and used within one week of dilution.

2.8.8 Hybridisation

The probes were denatured at 80°C for 2 min, cooled on ice and 10µl of probe was pipetted onto each prehybridised tissue section (being careful not to scratch the sections). A coverslip was placed over the sections avoiding air bubbles. The slides were placed coverslips upwards, in a hybridisation box containing tissue paper soaked in 50% v/v formamide, 5 x SSC. The box was then sealed with tape and placed into two water tight plastic bags. The water tight hybridisation box was placed in a water bath overnight at 55°C.

2.8.9 Post-hybridisation washes

The hybridisation box was removed from the waterbath and the bags were discarded. Forceps were used to transfer the slides to a slide rack which was then placed in a glass trough containing 5 x SSC, 10mM DTT for 30 min at 55°C. The coverslips usually fell off at this stage, if they had not, they were gently removed. The slides were transferred into HS wash (50% v/v formamide, 2X SSC, and 0.1M DTT added just before use) for 30 min at 65-68°C, NTE (0.5M NaCl, 10mM Tris, 5mM EDTA) wash for 2 x 5 min at 37°C, NTE wash with 20µg/ml RNase for 30 min at 37°C, NTE wash for 5 min at 37°C, HS wash for 30 min at 65-68°C. The remaining washes were performed at room temperature: 2 x SSC for 4 x 10 min, 0.1 x SSC for 4 x 5 min, 30%, 50%, 70%, and 90% v/v ethanol (each with 0.3M ammonium acetate) for 1 min each, 100% ethanol for 2 x 5 min and airdried.

Slides were exposed to x-ray film overnight at room temperature to give an estimation of the length of time required for exposure to emulsion.

2.8.10 Coating slides with emulsion

This procedure was carried out in a darkroom with a safety light on but at least 1m from the working area (the emulsion is extremely light sensitive). The LM-1 hypercoat emulsion (Amersham Life Science) was placed in a waterbath at 42°C for 15-20 min. Once the emulsion had melted, it was poured into a dipping chamber (Amersham Life Science). Slides were dipped vertically into the emulsion and left for 5 sec. The slide was then slowly withdrawn keeping it vertical and the emulsion was allowed to drain for 5 sec. The slide was allowed to stand nearly upright (approx 80° angle) allowing the emulsion to begin to set (about five min). Semi-set slides were placed in a slide rack and the rack was sealed into a light-tight box containing damp tissue paper (to avoid drying the emulsion too fast which leads to cracks). After ~3 hour the slides were transferred to a slide box containing silica gel, which was taped and wrapped in two layers of aluminium foil. The slides were exposed at 4°C for 1-8 weeks.

2.8.11 Development of emulsion-coated slides

Slides were allowed to reach room temperature before proceeding (approximately 3 hours). The slides were unwrapped in the dark room with the safe light on. It is critical for the developing that all solutions were within 2°C of each other (15-20°C is optimum; above this and the emulsion may begin to melt). The slides were placed in glass racks and immersed in troughs of developer for 4 min (1:4 dilution of Phenisol Developer to distilled water), Stop solution (0.5% v/v acetic acid) for 1 min, Fixer (30% w/v sodium thiosulphate) for 8 min after which the lights can be switched on, gentle running water for 15 min, followed by 2 x 15 min washes in distilled

water. The slides were carefully wiped to remove any remaining emulsion (especially on the back).

Solutions for counter staining must be within 2°C of the developing chemicals. The slides were counter stained using Giesma (1:20 dilution of Giesma R66 stock and Gurr's buffer pH6.5) for 15 min, running water for 15 min, 70%, 90% v/v and then 2 x 100% ethanol for 2 min each, followed by xylene for 2 x 2 min. The slides were air dried. Slides were then mounted using DPX and clean cover slips (any residual liquid on the slides crystallises in the DPX; therefore it is important that the slides are dry).

2.8.12 Photography of tissue *in situ* slides

Tissue *in situ* slides were examined using bright-field illumination on an Olympus BH-2 microscope fitted with neutral density filters 25 and 6, an Olympus BH2-SC swing out condenser and a Splan x4,x10 or x20 objective. An Olympus U-FT focusing telescope aided sharp focusing of the images which were recorded using an Olympus C-35AD-4 camera and Kodak 64T film.

Tissue *in situ* slides were also examined under dark-field illumination at the Western General Hospital HGU Unit with the help of Dr. Paul Perry. A Zeiss Axioplan 2 microscope with Zeiss Plan-neofluar x10 objective was used with an HBO 100 Mercury arc lamp. The images were captured using IP Lab Software. Three images are generated from this: a normal light field image, the silver grains reflected by dark field, and a composite image of the two whereby the silver grains are visualised in yellow and overlayed onto the coloured image from the light field.

2.8.13 Quantification of *in situ* results

Mammary gland sections were counted after *in situ* hybridisation to determine the percentage of the gland expressing. 10 fields of view at x 20 magnification were counted per section with at least two sections per animal being counted. Alveoli were scored as negative, -/+ (when some cells within an alveolus were positive while others were negative) or positive. Statistical analysis was carried out using GenStat with advice from a Statistician, Caroline McCorquodale. A non-parametric test (Mann-Whitney) was used to analyse the data from the different copy number lines. A Mann-Whitney test does not require the assumption of normally distributed observations in the two groups being compared.

2.9 RNA ANALYSIS

2.9.1 Preparation of total RNA from mouse tissue

All glassware, pipettes and solutions used for RNA analysis were treated with diethyl pyrocarbonate (DEPC) to destroy ribonucleases and gloves were changed frequently.

Tissue was frozen in liquid nitrogen immediately after dissection from the animal. Frozen tissue was homogenised in 2 ml RNAzol B (Biogenesis Ltd) for 30 sec using an ultra-homogeniser (Janke and Kunel). 2 x 1 ml aliquots were added to 2 ml screw-capped tubes containing 100µl chloroform. These were inverted repeatedly for 15 sec, put on ice for 15 min, and then microcentrifuged (13000 rpm, for 15 min at 4°C). The aqueous

phase was transferred to a fresh tube and the RNA was precipitated using an equal volume of isopropanol (4°C for 15 min). The samples were microcentrifuged (13000 rpm, for 15 min at 4°C) to pellet the RNA. The pellets were washed twice with 70% v/v ethanol (microcentrifuged 13000 rpm, for 8 min at 4°C between washes). The RNA was resuspended in 200µl of deionised formamide and left at 4°C until in solution. The RNA concentration was determined by measuring the absorbance at 260nm of a 0.01 dilution of each sample in TE buffer.

2.9.2 Electrophoresis of RNA, Northern Blotting and hybridisation

Messenger RNAs have secondary structures. Denaturing gels, therefore, are used for RNA electrophoresis.

10µg of RNA was added to 20µl sample buffer (3.5µl formaldehyde, 2µl 10 X MOPS, 4µl loading dye [30% Ficoll, 0.1M EDTA, 0.25% w/v bromophenol blue], 0.5µl 10mg/ml EtBr, 10µl distilled water) before loading onto a denaturing gel. The samples were usually electrophoresed overnight at 35V or at 80 V for 3 hours in denaturing gels (1.5% w/v agarose, 1 x MOPS, 12.5% v/v formaldehyde) using 1 x MOPS as the gel tank buffer. Northern blots were performed as described in section 2.5.11.1 except there was no pre-treatment of the gel before blotting onto Zetaprobe memebrane (Bio-Rad). Preparation of probes and hybridisation conditions were identical to those described in section 2.5.12.

2.10 QUANTIFICATION SOFTWARE

The quantification software used for both Southern blotting and SDS-PAGE gels was Quantity One (BioRad). Quantity One can be used to image and analyse data from phosphor imagers and gel documentation systems. Signals from biological samples are converted into digital data. A data object as displayed on the computer is composed of individual screen pixels. Each pixel has an X and Y co-ordinate, and a value Z which represents the signal intensity. Once an image has been captured, Quantity One Volume tools are used to quantify data (Quantity One User Guide 4, BioRad 2000).

Essentially a 'volume rectangle' is drawn around each data object. If the data represents a known standard its concentration is included. Global background is subtracted from the volumes by defining a background volume which corresponds to a region where there is no data, and where the average pixel intensity appears to be the same as the background intensity surrounding the data. A volume analysis report is generated once all data have been given a defined volume. If bands from standards with known concentrations are included on the gel / blot, the report produces a concentration value as calculated from the named standards. A volume regression curve is also generated from the standards, with unknown samples being placed on the curve.

CHAPTER THREE

PRODUCTION OF LOX TRANSGENIC LINES

3.1 INTRODUCTION

As discussed in Chapter one, observations of BLG transgene variegation had been found and high copy number arrays were thought to be involved (Dobie *et al* 1996). There was a need to determine what effect copy number in its own right had on BLG expression patterns, without the involvement of position effects. This project's aim was to specifically address the relationship between variegation and transgene copy number. With this in mind a construct was made which would allow manipulation of copy number at a constant genomic integration site. This was accomplished by the use of the Cre recombinase / Lox system providing a site-specific deletion within an animal (figure 3.1).

To investigate the relationship between high transgene copy number and variegation, high copy number lines were produced by pronuclear injection of a lox containing BLG transgene, and their expression profile analysed. Milk, RNA and DNA were all analysed to give an indication of variegation as discussed in this chapter. Animals from variegated lines were bred to a Cre recombinase line, or used to produce embryos for microinjection with Cre recombinase, and these lines were further analysed as discussed in Chapters 4 and 5.

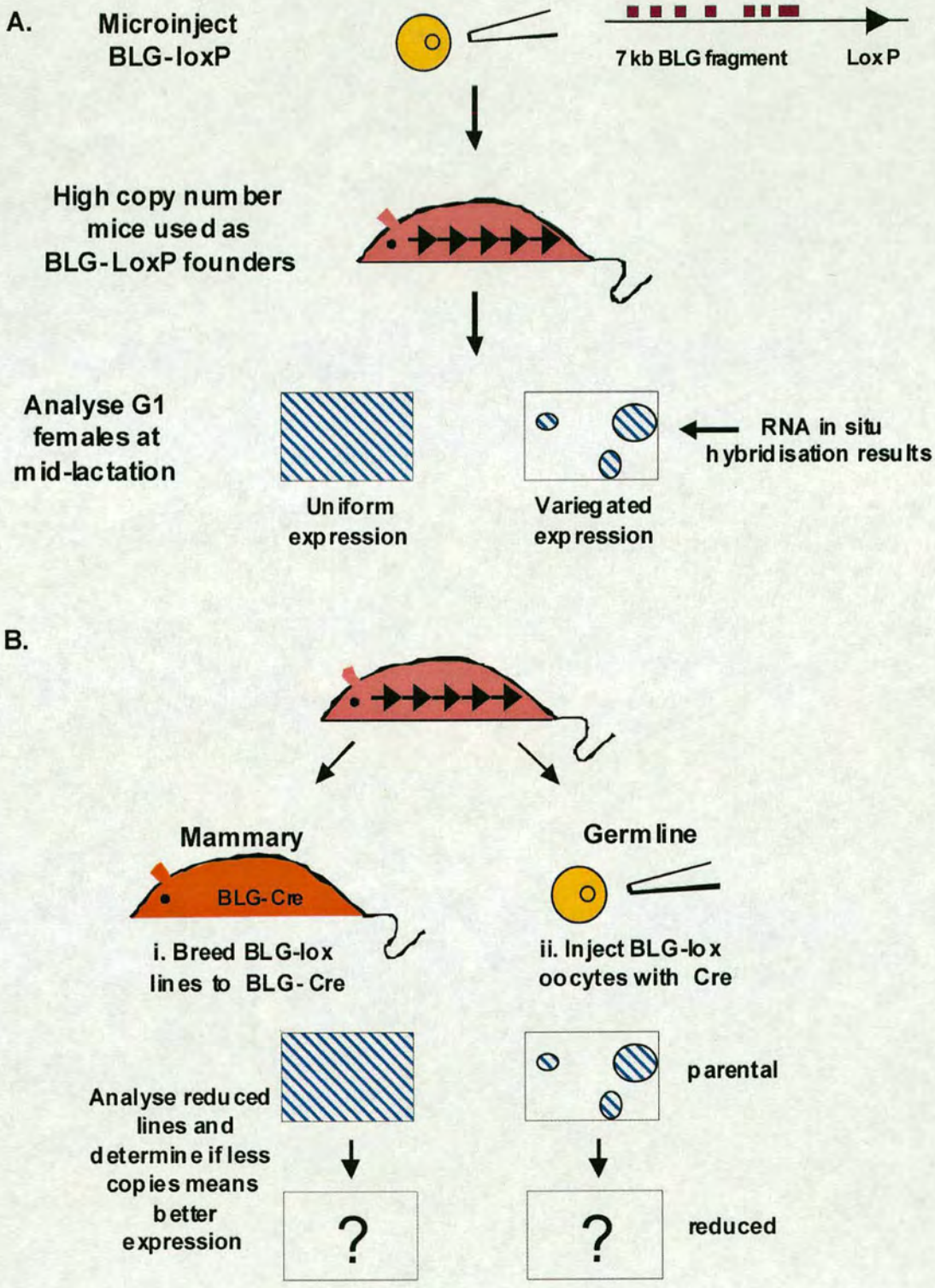


Figure 3.1. Strategy for site-specific recombination to alter the copy number of a transgenic array. **A.** Production of high copy number lines. **B.** The reduction of copy number would be achieved by two different routes allowing comparisons between different developmental time points.

3.2 RESULTS

3.2.1 Analysis of BLG-loxP construct

The strategy of manipulating the transgene copy number required a number of prerequisites. The transgene used must contain a loxP site to allow Cre mediated recombination to occur, and the transgenes needed to be integrated in such a manner that the loxP sites were in the same orientation. If the transgenes had integrated in a head to tail array then after the expression of Cre recombinase, the transgenes between adjacent loxP sites could be deleted, thereby causing a reduction in copy number (figure 3.2).

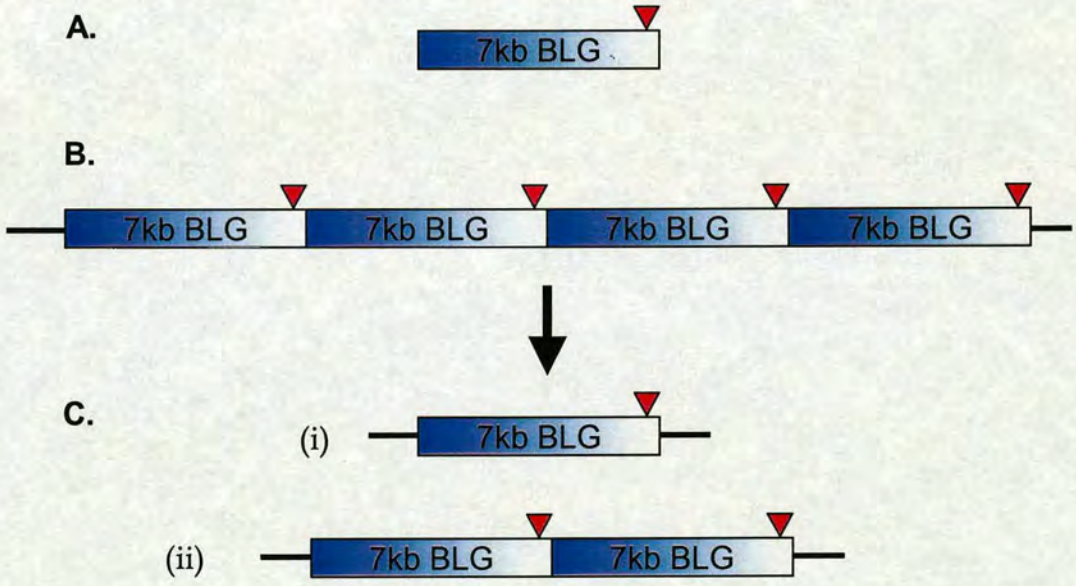


Figure 3.2. Requirements for reduction in copy number strategy. (A) BLG-loxP transgene which would be a target for Cre recombinase. Red triangle represents a loxP site, blue box represents the BLG transgene. (B) Head to tail transgenic array integrated into the genome (black line). (C) This arrangement would produce a reduction in copy number after Cre recombination. The reduction could be to (i) a single copy or (ii) more copies could be retained depending on the efficiency of the Cre mediated recombination.

The BLG-loxP construct involved had one loxP site inserted into the 3′ end of the transgene (a gift from Dr. Douglas Strathdee; Roslin Institute). A loxP oligonucleotide had been cloned into site 10320 of the pPolyΔDp BLG construct. The oligonucleotide was designed such that once integrated it would provide a fourth BspHI restriction site, producing a new band of 1.4kb. A diagnostic restriction digest confirmed that one loxP site had been introduced at the 3′ end of the BLG construct (figure 3.3).

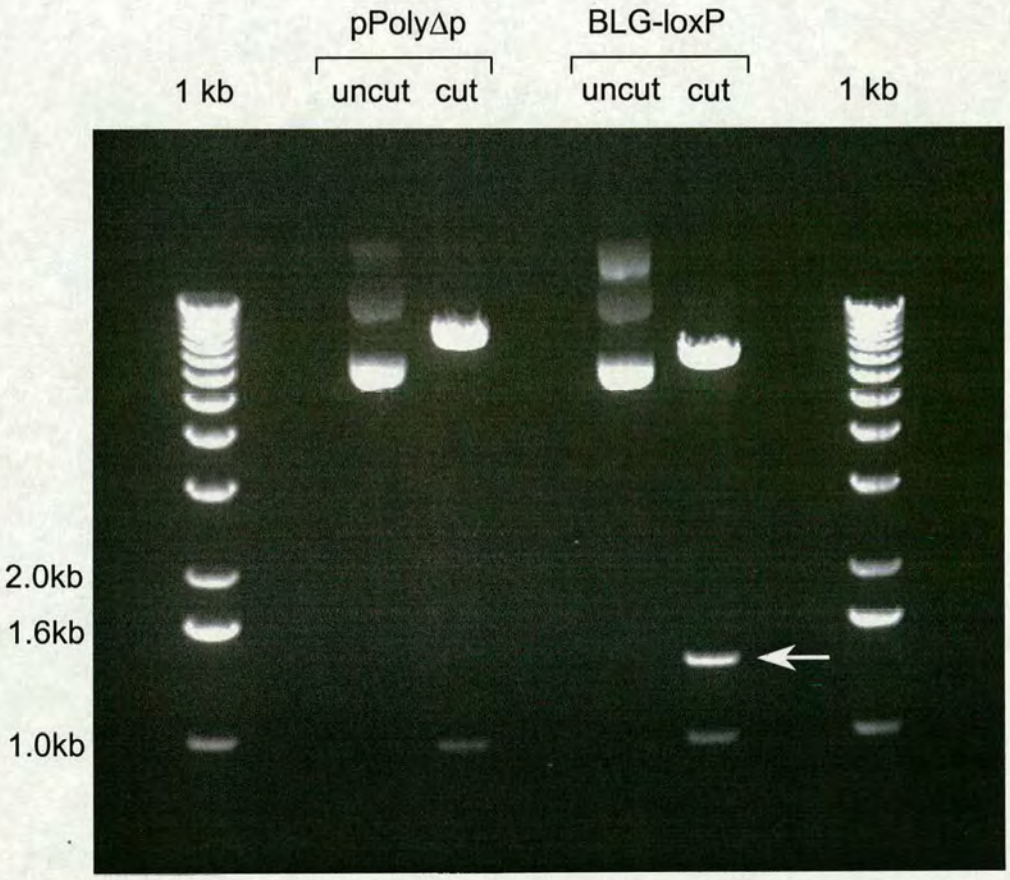


Figure 3.3. Diagnostic restriction digest of BLG-loxP construct. To determine the number of lox oligonucleotides integrated, the BLG-loxP construct DNA was digested with or without BspHI for 2 hours and compared with the original construct pPolyΔDp. The digests were loaded on a 1% agarose gel. The predicted band of 1470 bp was present (indicated by arrow). 1kb ladder from GibcoBRL with stated fragment sizes. The BspHI digestion also produced a 105bp fragment from both constructs which is not visible in this figure.

Cre recombinase is very site specific and even single base pair changes at the recognition site can prevent the reaction. To ensure that the introduction of the loxP site had not led to a point mutation within this area DNA sequencing was carried out (figure 3.4). This confirmed that only one loxP site had integrated, rather than two loxP sites at the same location and that the loxP site had integrated without mutation.

tgatctcatg ataacttcgt ataatgtatg ctatacgaag
ttatgatcag gatcaggacg atgggaagac tgtgagcacc
ctgaaggcgg ggccaccccc gagtgtctgg tgactcatag
gtgaaggctg gctccctgaa gcccctgagg cctggggagg
agg

Figure 3.4. Sequence data from BLG-loxP construct. Underlined letters indicate the loxP site which has integrated into the BLG sequence with no mutation. The sequence starts at 10325 bp and finishes at 10212bp using the numbering from BLG sequence x12817.

3.2.2 Microinjection of BLG-loxP.

Once the integrity of the loxP site was established, DNA was prepared for microinjection. The BLG-loxP construct was digested with XhoI to remove vector sequence before being used for pronuclear microinjection, as vector sequences have previously been shown to inhibit transgene expression (Townes *et al* 1985).

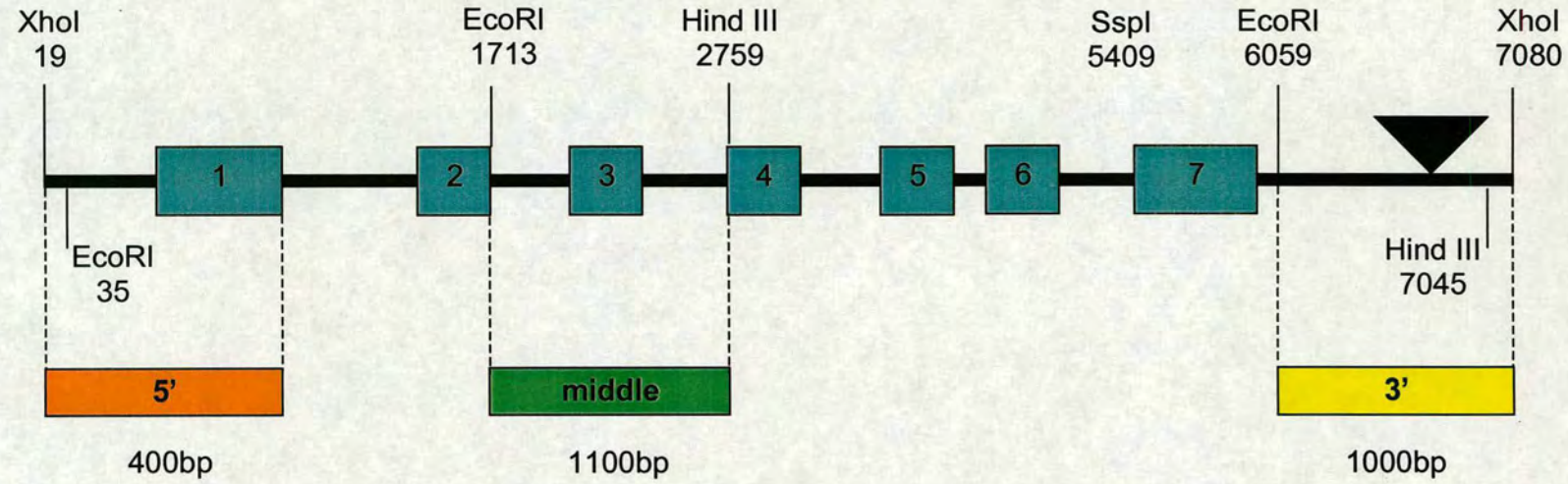


Figure 3.5 Probes used for Southern blotting. The plasmids from which the fragments were derived from are listed in Appendix One. The figures given for restriction sites are from BLGΔDp sequence. The diagram is not to scale.

3.2.3 Analysis of transgenic mice.

Having established that the BLG-loxP construct contained an intact loxP site, it was now necessary to analyse founder animals for the structure and copy number of the integrated BLG-loxP transgene. For the project strategy to be successful, the transgenes needed to be integrated in a head to tail array. Of the 95 founders produced by microinjection, 14 were deemed positive by Southern blot (probes used for Southern blotting are found in figure 3.5). Five high copy number founders were then used to establish transgenic lines. These lines were maintained in a hemizygous state by testing with PCR/Southern Blotting.

3.2.4 Analysis of transgene copy number

Copy number estimates of the five established lines were made. The results were obtained by comparing the intensity of internal transgene fragments on Southern blots of G1 liver DNA to copy control standards from dilutions of the original 7kb microinjection fragment. The copy control standards were calculated using the following assumptions: that the haploid content of the mouse genome is 3×10^9 bp and secondly that the copy controls will be used in a 10μg digestion of genomic DNA (Camper 1987). Since the mice in this study are heterozygous:

$$\frac{\text{Mass of transgene DNA}}{\text{5 micrograms genomic DNA}} = \frac{\text{N bp transgene DNA}}{3 \times 10^9 \text{ bp genomic DNA}}$$

e.g. for BLG-loxP:

$$\frac{\text{Mass of transgene DNA}}{\text{5 micrograms genomic DNA}} = \frac{\text{7080 bp transgene DNA}}{3 \times 10^9 \text{ bp genomic DNA}}$$

e.g. for BLG-loxP transgene: 11.8 picograms is equivalent to one copy.

Once the amount of transgene DNA corresponding to one, 10, 20 copies etc. had been determined, this amount was added to 10μg of non-transgenic DNA and digested as appropriate. The probed blots (figure 3.6) were scanned and quantified by Quantity One software. The copy controls were used as known standards to produce a standard curve which was then used to estimate the unknown transgenic copy number (see 2.10). From these blots the copy numbers were estimated as described in Table 3.1.

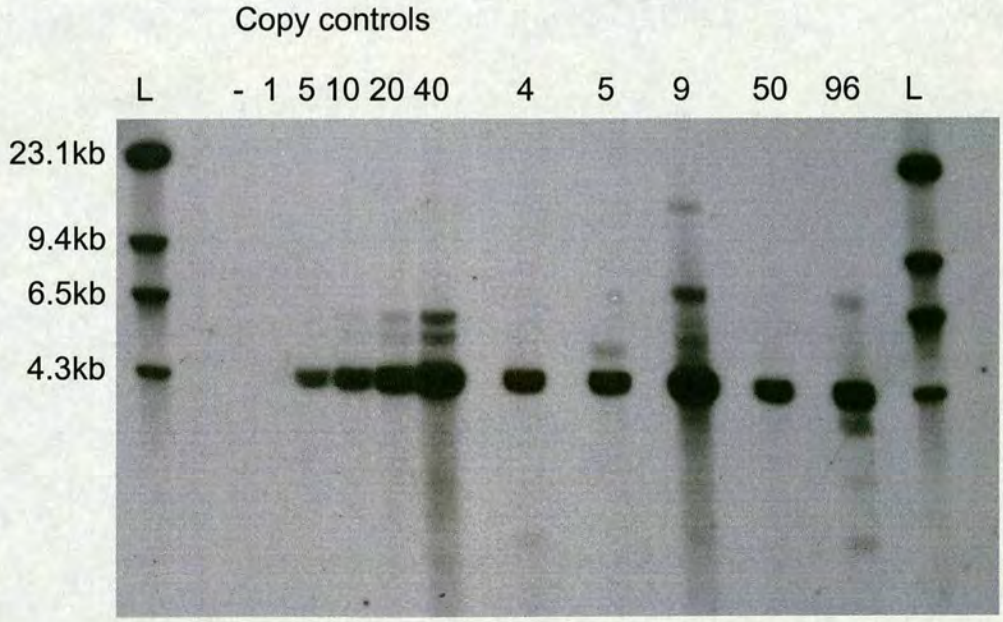


Figure 3.6. Analysis of lox transgene copy number by Southern blot. 10μg genomic liver DNA was digested with *Eco*RI, electrophoresed on a 0.8% agarose gel and probed with a 1.1 kb BLG fragment. L indicates a λ HindIII ladder with stated fragment sizes which was used for size estimation. – indicates 10μg of non-transgenic DNA. Copy controls were dilutions of the original microinjected 7kb BLG-loxP fragment added to 10μg of non-transgenic DNA.

In figure 3.6 the *Eco*RI digest produced a fragment of 4.3kb in all of the transgenic lines which corresponded to an internal 4.3kb fragment of the BLG-loxP transgene (see figure 3.5). Lox 50 shows no other transgenic fragments. However, in four of the transgenic samples other fragments were present.

In the lox 4 sample a faint fragment (approx 1.5kb) can be seen. This may represent a truncated transgene which has lost some of the sequence that corresponds to this probe, producing a fragment smaller than the expected 4.3kb.

In the lox 5 sample a larger fragment (approx 4.6kb) is present at lower intensity than the internal 4.3kb fragment. This may represent another truncated fragment which contains the 4.3kb EcoRI fragment and additional sequence around it.

In lox 9 multiple fragments are present. Given that one of these (approx 5.2kb) corresponds to the partial digestion fragments of the copy control, these may represent poorly digested bands or transgenes that have lost one of the *Eco*RI sites. There is also a band at approx. 7kb which would correspond with the undigested transgene length. The largest fragment (>10kb) in lox 9 may represent the junction fragment of this transgenic array.

In lox 96 the 7kb fragment may represent undigested transgenic DNA, with the smaller fragments (approx 3.3kb, 2.4kb and 1.3kb) representing truncated regions of transgene or junction fragments containing only a part of the 4.3kb band.

The method of comparing copy controls against internal transgene fragments is however susceptible to a number of errors. Accurate results depend on a number of factors: the ability to accurately determine the concentration of the DNA, the ability to accurately dilute the copy control fragment and lastly accurate pipetting of genomic DNA. The ability to accurately 'box' a data point using quantification software may also lead to errors in the estimation of copy number. The copy controls are also 'external', as they are not derived from the same biological sample as the unknowns and are electrophoresed on different lanes of the gel.

3.2.5 Estimation of copy number by junction fragment comparisons and dilution blots

An alternative method was used to estimate transgene copy number of the arrays, by comparing the transgene and its junction fragments. This represents an internal control as both are derived from the same sample and are electrophoresed on the same lane of a gel. An enzyme with a unique restriction site within the transgene was used to digest the DNA. If the array has more than one copy it produces a repeat length fragment and junction fragments. Comparisons can be made between the intensity of the internal repeat fragment and the junction fragment. As the 1.1kb probe can only hybridise to one site per transgene copy, the copy number of the transgene repeat can be determined by its intensity, relative to the junction fragment, which is used as a one copy control. The junction fragment between the transgene array and the genomic sequence is unique and is therefore only represented once per genome.

Southern blots were exposed to a Phosphoimager screen (BioRad) for quantification. Phosphoimager screens have a larger linear range than X-ray film (5 orders of magnitude; BioRad) and can also be used to determine if any bands have become saturated allowing another exposure to be used. This means that they can be used to compare bands of quite different intensities from the same gel as both will fall within the linear range of the phosphoimager screen. Quantification of copy number based on band intensity assumes that the probe hybridises equally to the transgenic sequence in both bands, and that it is not affected by fragment length or other genomic sequence. It also assumes that the junction copy is only present once per genome. Quantification may be hampered by errors in DNA concentration and dilutions, accurate boxing of data points and the sensitivity of the phosphor screen.

Dilution blots provide another method of copy number estimation. The idea is to dilute the repeat fragment until the signal from it matches that of the junction fragments when hybridised with the same probe. The dilution factor required to bring the internal fragment to the same intensity as the 6kb junction fragment gives an estimation of the copy number (figure 3.7).

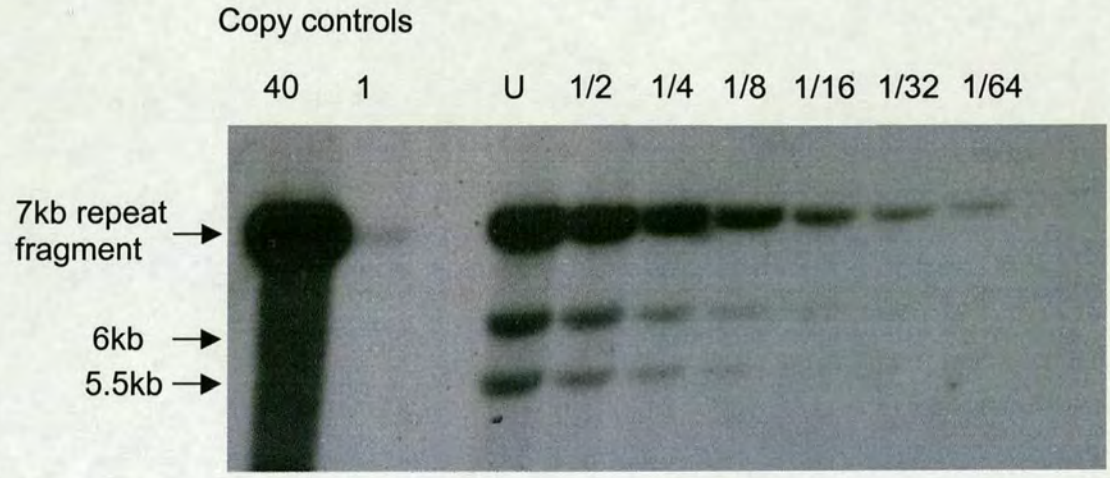


Figure 3.7. Dilution Southern blots to determine lox copy number. Transgenic DNA was either left undiluted (U) or diluted (numbers give dilutions) with non-transgenic DNA to give 10μg which was digested with *Ssp*I, electrophoresed on a 0.8% gel and probed with a 1.1kb BLG fragment. Copy controls were dilutions of the original microinjected 7kb BLG-loxP fragment added to 10μg of non-transgenic DNA.

From these measurements the lines lox 4, 5 and 9 were determined to have copy numbers as described in table 3.1. The two methods used gave different estimates (especially for lox 5 or 9), which emphasises how difficult it is to precisely estimate copy number. An alternative method is to use the length of the transgenic array as an estimate of the number of copies. Pulse Field Gel Electrophoresis (PFGE) of a digest which does not cut within the transgenic array will give a transgenic array fragment of a certain length, which can be used to determine the number of copies present i.e. 18 copies of a 7kb transgene would be 128kb in total length. This method avoids any difficulties in comparing intensities of bands as discussed previously.

Line	*Estimated copy number by comparison to copy controls*	*Estimated copy number by comparison to junction fragments*
Lox 4	16	18
Lox 5	13	26
Lox 9	21	35
Lox 50	11	NA
Lox 96	40	NA

Table 3.1. Estimates of lox transgene copy number by different methods. Copy control comparisons are from figure 3.6. *Ssp*I digestions were used to determine copy number by finding the nearest dilution in intensity to the undiluted junction fragment. The ratio of repeat fragment counts/junction fragment counts was used to determine the number of repeats actually found in the array. Dilution blots were only carried out for lox 4, 5 and 9.

In figure 3.7 there is the 7kb repeat length fragment and two other fragments: a 6kb fragment and a 5.5kb fragment. The 6kb was interpreted as the 5′ junction fragment corresponding to genomic sequence and the first 5.4kb of the transgene up to the *Ssp*I site. However, the 5.5kb fragment could not represent a 3′ junction as the probe used in this blot would not hybridise to sequence 3′ of the *Ssp*I site. Furthermore as discussed in Chapter 5 this 5.5kb fragment is absent after Cre mediated recombination. The site-specific recombination should result in the same junction fragments at the end of the transgenic array as before the introduction of Cre recombinase. Taken together, it is possible that the 5.5kb *Ssp*I fragment therefore, represents an internal truncated fragment of the transgene which is removed along with other copies of the transgene upon the introduction of the Cre recombinase.

3.2.6 Analysis of the transgenic array structure

Critical to the strategy for the redution of copy number was the use of of transgenic lines that contained head to tail arrays. If the arrays are not head to tail, then the loxP sites are in the wrong orientation to produce a Cre mediated deletion between them. To determine if the lox transgenic lines had integrated as head to tail arrays the DNA was restriction digested with *Ssp*I, an enzyme that cuts once within the transgene. If the array comprises solely of head to tail repeats it will produce a 7kb fragment, while 10.5kb or 5kb fragments would be produced if there were head to head, or tail to tail arrays (figure 3.8).

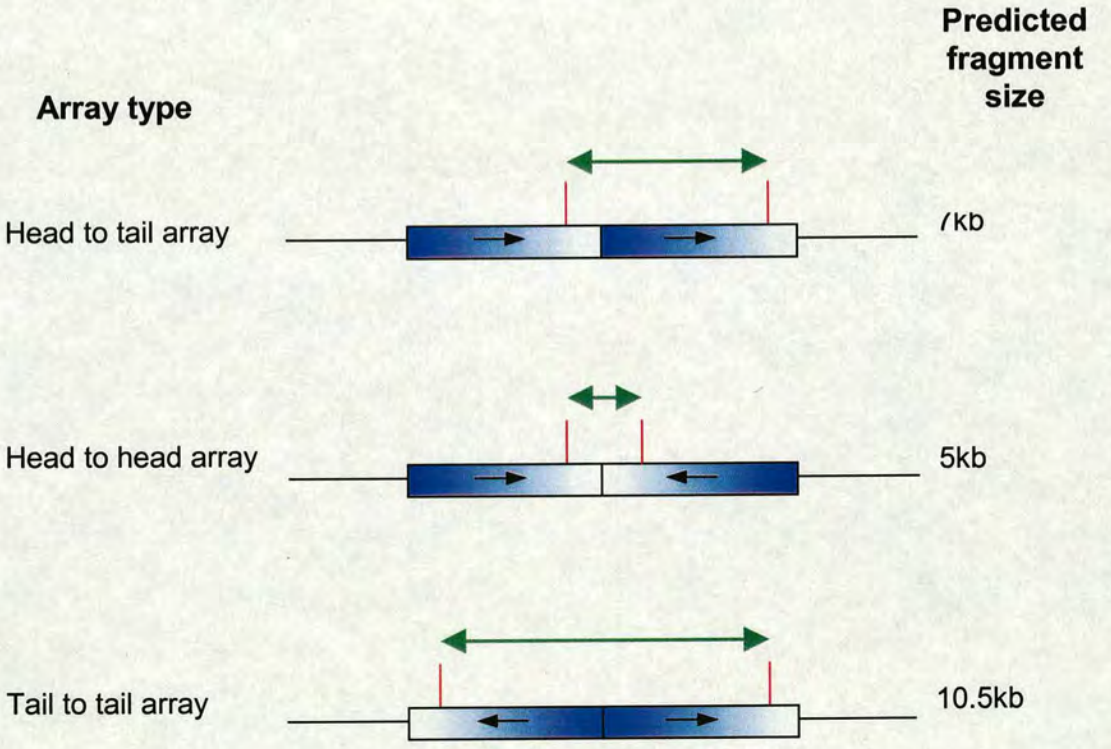


Figure 3.8. Predicted *Ssp*I fragments from different transgenic array structures. The red lines represent where the *Ssp*I site is on the transgene. The green arrow represents the repeat fragments expected from each type of array.

As seen from figure 3.9, lox 4, 5 and 50 contained head to tail repeats giving only the predicted 7kb repeat fragment and junction fragments. Lox 9 had the 7kb repeat fragment from head to tail arrays, but it also contained three other transgenic fragment indicating that at the integration sites there may be other segments of DNA integrated between the transgene and the ends of the genomic sequence or rearrangements of the transgene. Lox 96 has a complex structure, which may indicate that it contains two closely associated but separate integration events. Therefore this line was not used for reduction events due to its complex array. This blot could not determine if any head to head transgenes existed due to the location of the probe relative to the restriction site for SspI. No individual probe in this study could hybridise to all three arrangements.

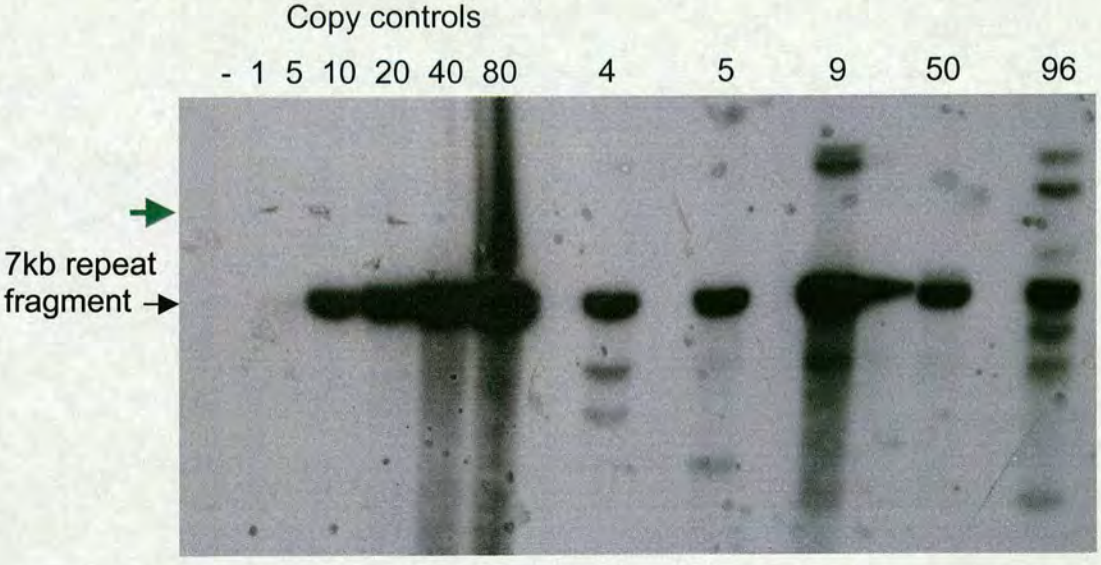


Figure 3.9. Analysis of transgene array structure by Southern blot. 10μg transgenic genomic liver DNA was digested with *Ssp*I, electrophoresed on a 0.8% gel and probed with a 1.1kb BLG fragment. SspI cuts once in the transgene producing a 7kb fragment if there are head to tail arrays, a 10.5kb fragment from a head to head array and a 5kb array from a tail to tail array. However, the probe used would only identify head to tail (black arrow) and tail to tail (green arrow indicates where this band would occur). Copy controls were dilutions of the original microinjected 7kb BLG-loxP fragment added to 10μg of non-transgenic DNA.

3.2.7 Expression Analysis

3.2.7.1 *In situ* hybridisation

To determine the cellular pattern of the transgene, *in situ* hybridisation studies using ^{35}S labelled BLG and β-casein mRNA probes were carried out on abutting sections from mammary tissue collected at mid-lactation from all five BLG-loxP lines.

The structure of the mid lactation mammary gland tissue is shown in figure 3.10. An alveolus is composed of a single layer of secretory epithelial cells surrounding a lumen or alveolar space. The secretory epithelial cells are surrounded by a basket network of myoepithelial cells, which in turn are surrounded by a basement membrane (made of connective tissue proteins). Blood vessels and capillaries run through the stroma (inter-alveolar spaces). The stroma also contains fibroblasts, leukocytes (plasma cells) and other connective tissue cells. Secretory epithelial cells secrete milk proteins into the alveolar space. Myoepithelial cells on the outside of the alveolus contract to squeeze the milk out into the ducts and finally to the nipple (Mepham 1987). Sections of the mammary gland at mid-lactation produce cross-sections of alveoli and the ductal system (figure 3.10). The sections were stained with either Giesma or methylene blue, both of which give a blue counterstain to the cells with darker staining of the nucleus. The secretory epithelial cells have a simple 'cuboidal' shape. The slender myoepithelial cells surround the secretory epithelial cells and have an elongated shape. They are difficult to visualise with general staining methods at mid-lactation but can be seen by silver staining on goat mammary gland (Mepham 1987).

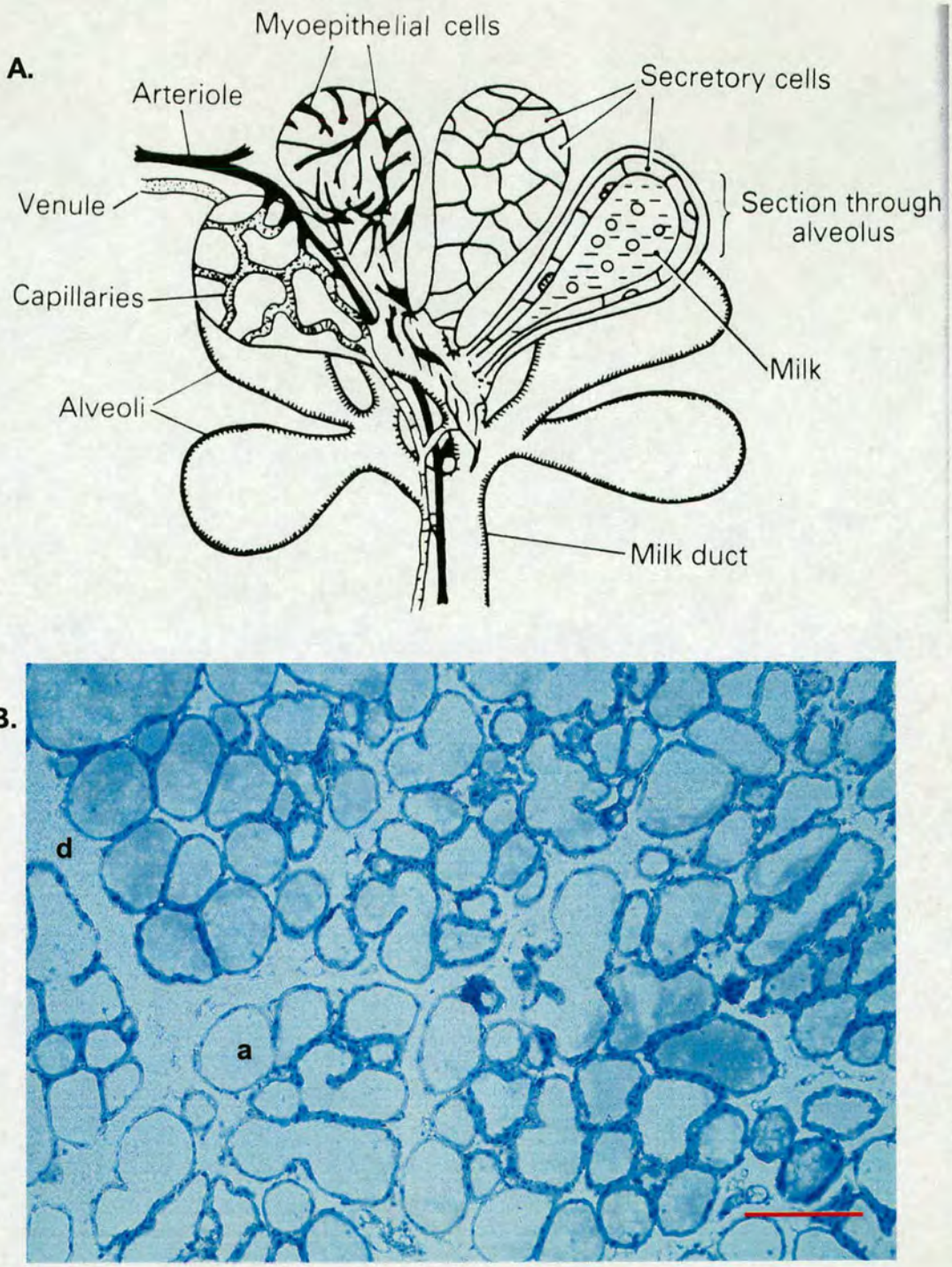


Figure 3.10. Morphology of the mammary gland. (A) Diagrammatic representation of a cluster of mammary alveoli (adapted from Mepham, 1987), (B) Section taken through a mid-lactation mouse mammary gland, probed with hybridisation buffer only. Photograph taken after 2 week exposure using a x10 objective. Bar represents 300μm. **a** indicates an alveolus, **d** indicates the ductal system.

Different markers can be used to differentiate between the various cell types found within the mammary gland. Antibodies to a variety of cell markers now exist allowing an investigator to distinguish between the cell types. Myoepithelial cells are contractile elements showing a combined epithelial and smooth muscle phenotype. Immunohistochemical markers for this cell type include smooth muscle actin and cytokeratin 14 (Sapino *et al* 1990). Different cytokeratins are found on different cell types: cytokeratin 14 is found on myoepithelial cells, while cytokeratin 8 represents secretory epithelial cells. Secretory epithelial cells can also be identified using beta-casein and alpha-lactalbumin (Sapino *et al* 1990). A good resourse for histology of the mammary gland is found at http://mammary.nih.gov/

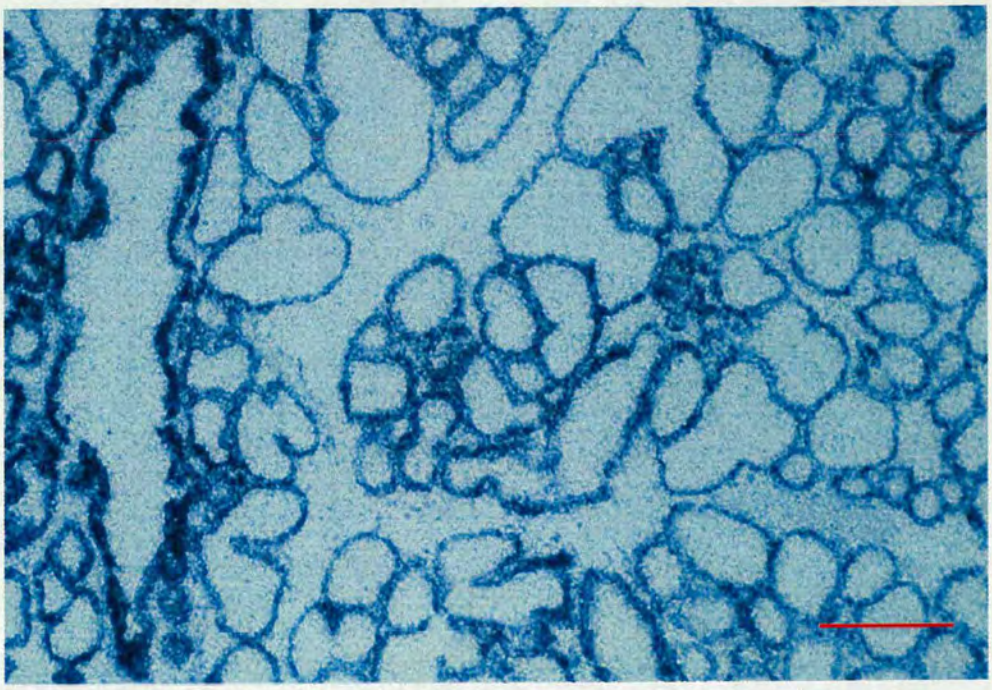
The mammary gland is a tissue composed of different cell types with the main constitutents being secretory epithelial cells and myoepithelial cells. A small number of studies have been carried out to determine the percentage of the gland represented by each type. However, it must be kept in mind that the mammary gland is an 'active' tissue, proliferating and involuting due to the varied influences of hormones and suckling, therefore the exact number of each cell type may vary depending on the different developmental stages it is sampled at. In one early study the DNA content of different cell types were estimated in C3H/Crgl/2 mice (Nicoll and Tucker 1965). It estimated that 89% of 10 day lactating mammary gland DNA was from mammary parenchyma. This figure would include both the secretory and myoepithelial cells. Subsequent studies have used immunohistochemistry to look at the percentage of cells and the proliferation of different cell types during development (Sapino *et al* 1990). In lactating mice secretory epithelial cells represented 70% of the gland, while myoepithelial cells represented 30% (Sapino *et al* 1990). In virgin mice

primed with estrogen and progesterone to mimic the changes to the mammary gland in early pregnancy, secretory epithelial cells represented 71% of the gland, myoepithelial cells represented 21% of the gland and other cells represented 8% (Sapino *et al* 1993). These few studies that distinguish between myoepithelial and secretory epithelial cells suggests that approximately 70% of the mammary gland is comprised of secretory epithelial cells.

In situ hybridisation was examined by both light and dark field microscopy. In light field images the ^{35}S labelled RNA when hybridised to target RNA in the cell causes a change in the silver grains of the photographic emulsion. These then appear as a dark black stain against the blue counterstain. If no RNA is present all that is seen is the counterstain. Under dark field microscopy only reflective particles are visible (i.e. only the the silver grains that have reacted to the ^{35}S labelled RNA) and these give a fuzzy silver appearance. When imaging software is applied, a dark field image can be overlayed onto a light field, allowing the silver grains to be visualised as a yellow 'paint'.

Non-transgenic mammary gland sections showed the expected expression of the endogenous gene β-casein (figure 3.11) while BLG mRNA was not detected even under dark field (figure 3.12). This demonstrated that the BLG mRNA probe did not hybridise in a non-specific manner to the tissue. Sheep mammary gland was used as a positive control and showed that as an endogenous gene the BLG mRNA is uniformly expressed, while the mouse β-casein probe failed to cross-hybridise to the sheep tissue (figure 3.13).

A.



B.

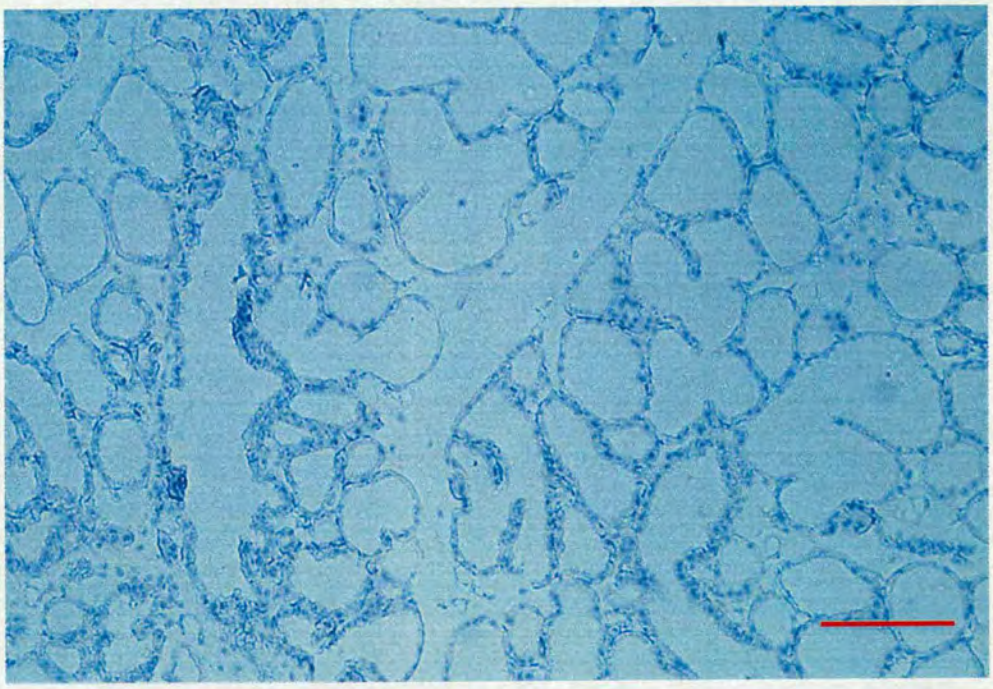


Figure 3.11. *In situ* hybridisation analysis of mRNA expression patterns in non-transgenic mice. (A) β-casein, (B) BLG. Photographs were taken after 2 week exposures using a x10 objective. Bar represents 300μm.

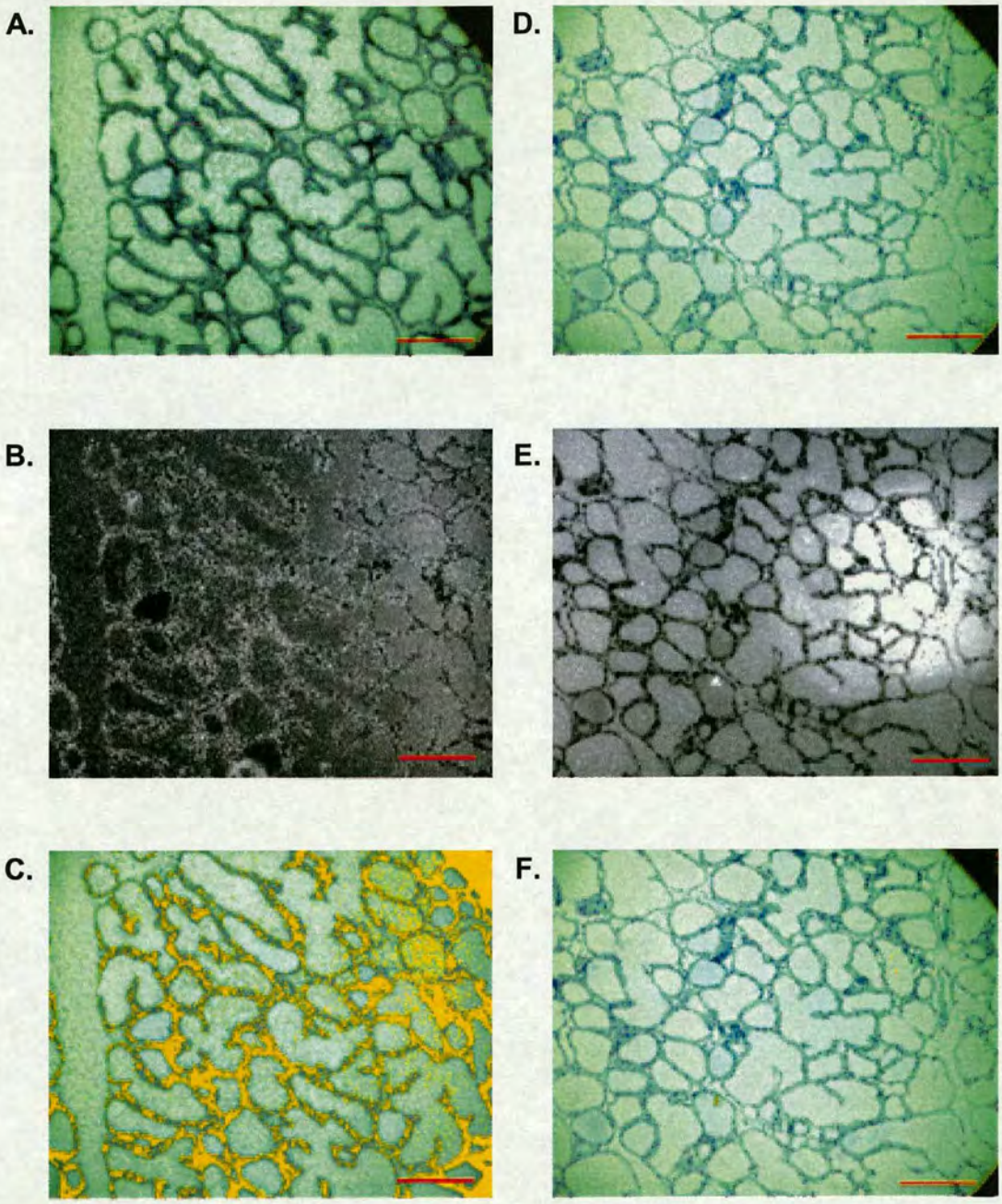


Figure 3.12. Dark field imaging of *in situ* hybridisation of non-transgenic tissue. (A,B,C) probed with β-casein. (D,E,F) probed with BLG. Photographs were taken after 2 weeks exposures using a x10 objective. (A) and (D) are light field images. (B) and (D) are dark field images. (C) and (F) are composite images of both light and dark field with the silver grains being represented with a yellow paint. Bar represents 300μm.

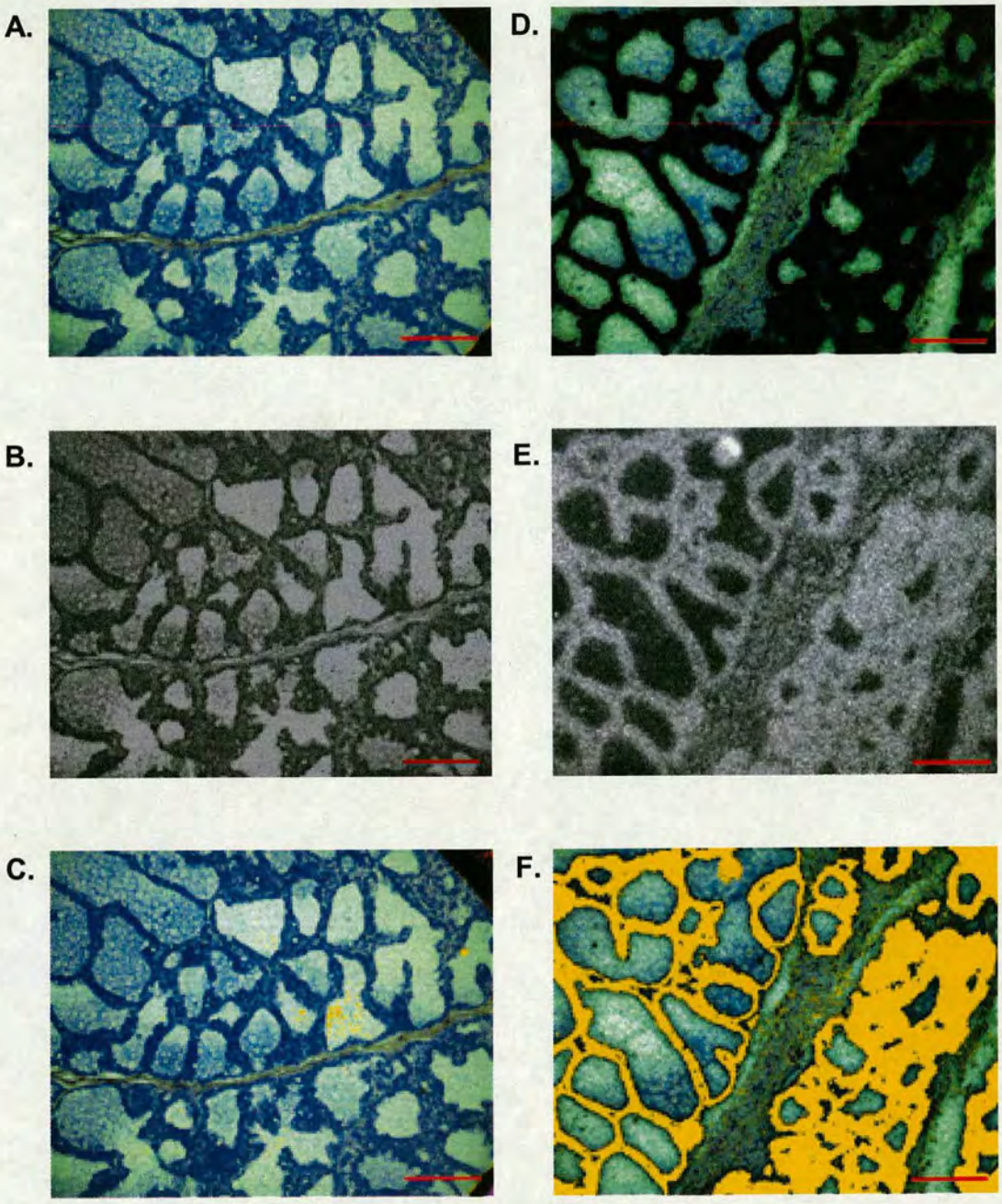


Figure 3.13. Dark field imaging of *in situ* hybridisation of sheep tissue. (A,B,C) probed with β-casein. (D,E,F) probed with BLG. Photographs were taken after 2 weeks exposures using a x10 objective. (A) and (D) are light field images. (B) and (D) are dark field images. (C) and (F) are composite images of both light and dark field with the silver grains being represented with a yellow paint. Bar represents 300μm.

Sections from all five lines probed with control BLG and β-casein sense probes had no signal associated with them by light field, although some did show very weak signal by dark field in inter-alveoli areas (figure 3.16, 3.19, 3.22, 3.25, 3.28).

β-casein mRNA was detected in a uniform manner throughout the secretory epithelial cells in all transgenic lox animals analysed. This demonstrates that essentially all the secretory epithelial cells had the capacity to express a milk protein gene.

Sections from lox 4 animals indicate that the BLG mRNA is present in a variegated manner, where discreet patches of cells express the transgene beside areas of non expressing cells. These distinct patches of expression may indicate that the variegated patterns seen are clonal in origin. However, some alveoli contained both expressing and non expressing cells (figure 3.14; a mixed alveolus is noted with a red arrow). In dark field imaging (figure 3.15, panel E) only some areas show reflective silver grains while the rest is dark.

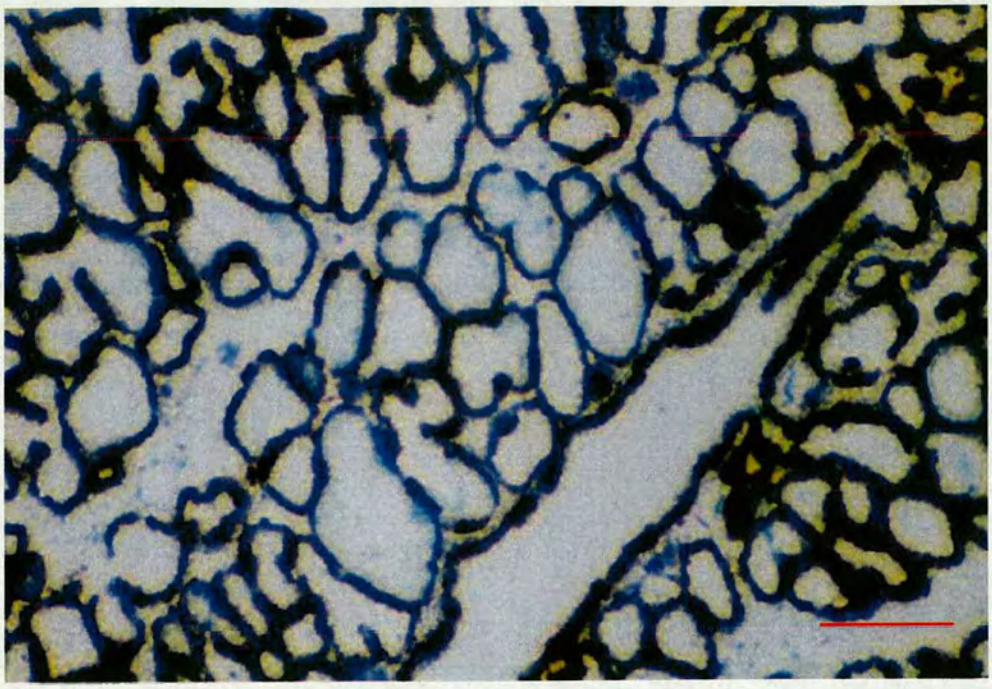
Lox 5 showed a uniform pattern of BLG expression throughout the mammary secretory epithelial cells (figures 3.17 and 3.18). Although darkly stained with methylene blue (figure 3.19) lox 5 sense probes showed no hybridisation indicated by the lack of a yellow overlay in panel C.

Sections from lox 9 animals showed that the BLG transgene mRNA was only present in discreet areas of the gland and was not uniformally expressed. Similarly to lox 4 animals some alveoli in lox 9 animals contained both expressing and non expressing cells (figure 3.20; red arrow).

Figure 3.21 shows that even by dark field imaging certain sections were completely silent for BLG expression (bottom left of panel E and F). Some background staining can be seen in ductal areas of lox 9 using a sense probe (figure 3.22).

Sections from lox 50 and lox 96 animals indicated that both these lines were uniform expressors of the BLG transgene mRNA (lox 50: figures 3.23 and 3.24, lox 96: figures 3.26 and 3.27). Because the photographic emulsion is at a slightly different focal plain than the tissue section, where there was high expression it was difficult to produce sharp images by normal light microscopy (i.e. 3.23). However, both showed uniform expression by dark field imaging (figures 3.24 and 3.27). Both these tissues showed no background with a sense probe (figures 3.25 and 3.28).

A.



B.

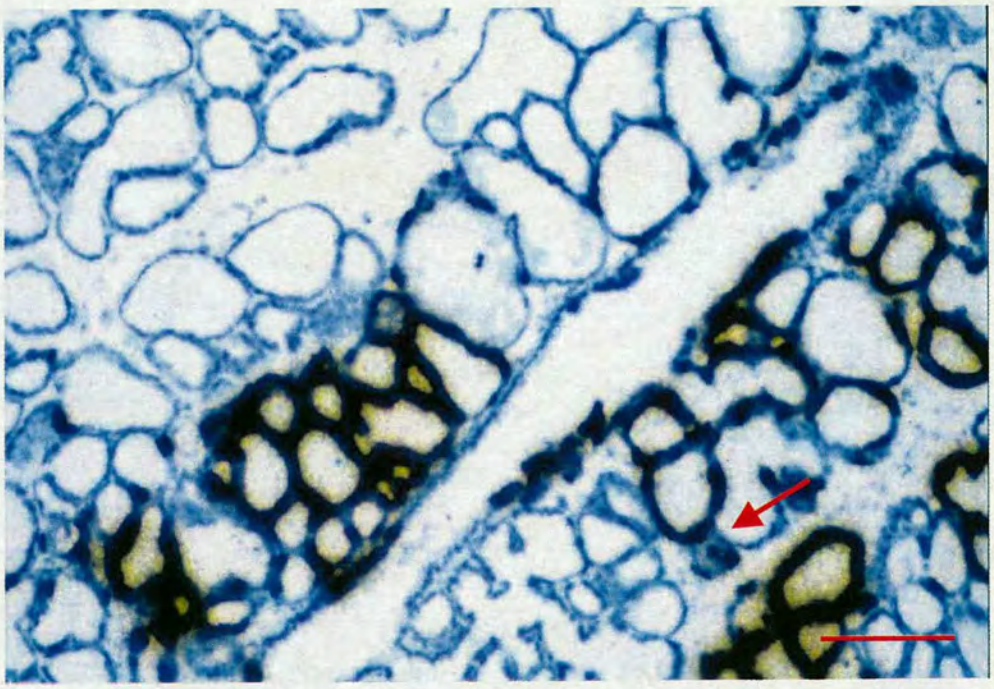


Figure 3.14. *In situ* hybridisation analysis of mRNA expression patterns in line BLG-lox 4.

(A) β-casein, (B) BLG. Photographs were taken after 2 week exposures using a x10 objective. Bar represents 300μm. Red arrow indicates an alveolus with partial expression.

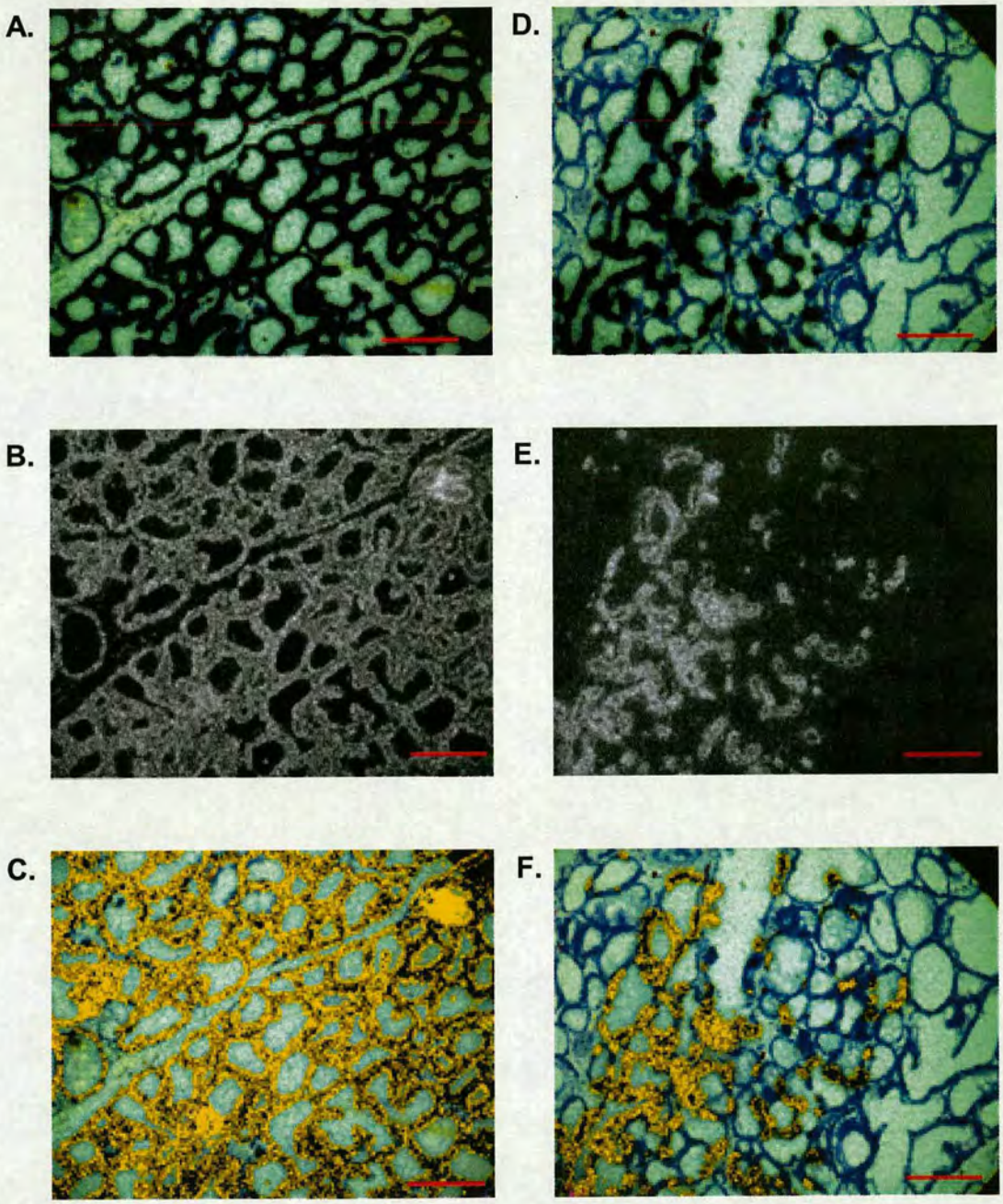


Figure 3.15. Dark field imaging of *in situ* hybridisation of Lox 4. (A,B,C) probed with β-casein. (D,E,F) probed with BLG. Photographs were taken after 2 weeks exposures using a x10 objective. (A) and (D) are light field images. (B) and (D) are dark field images. (C) and (F) are composite images of both light and dark field with the reflective silver grains being represented with a yellow paint. Bar represents 300μm.

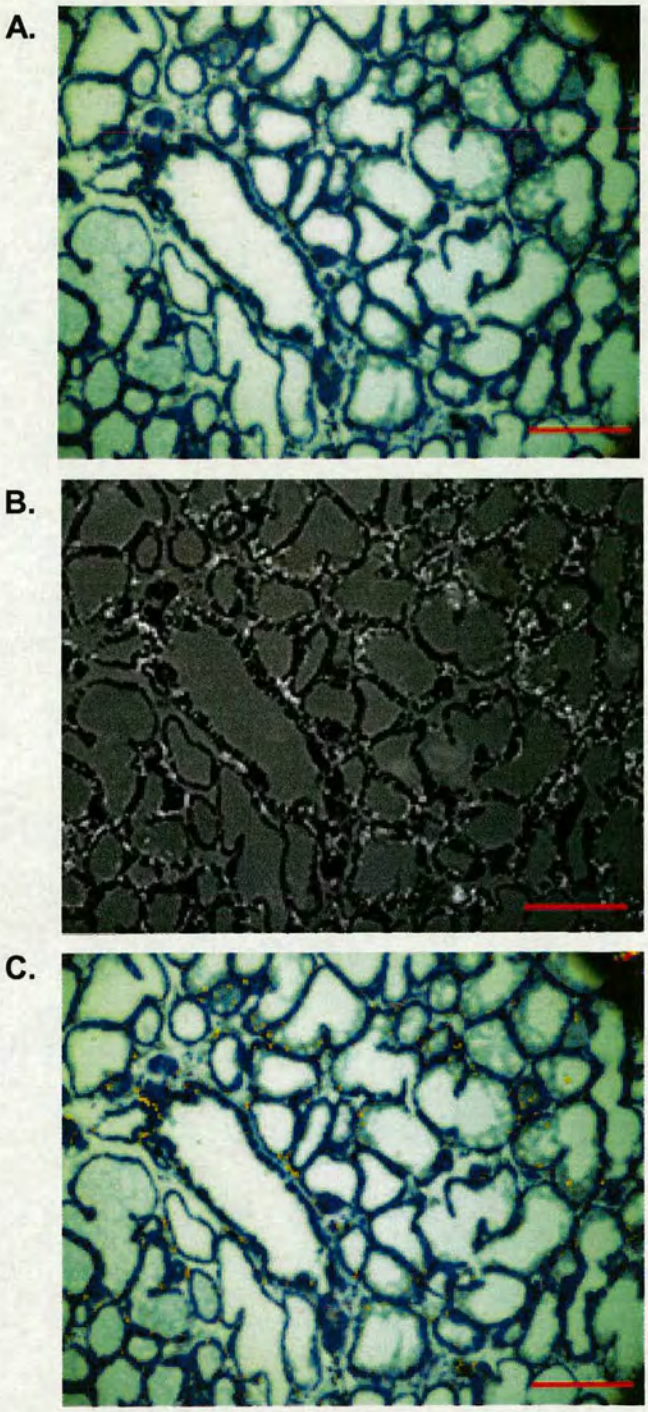
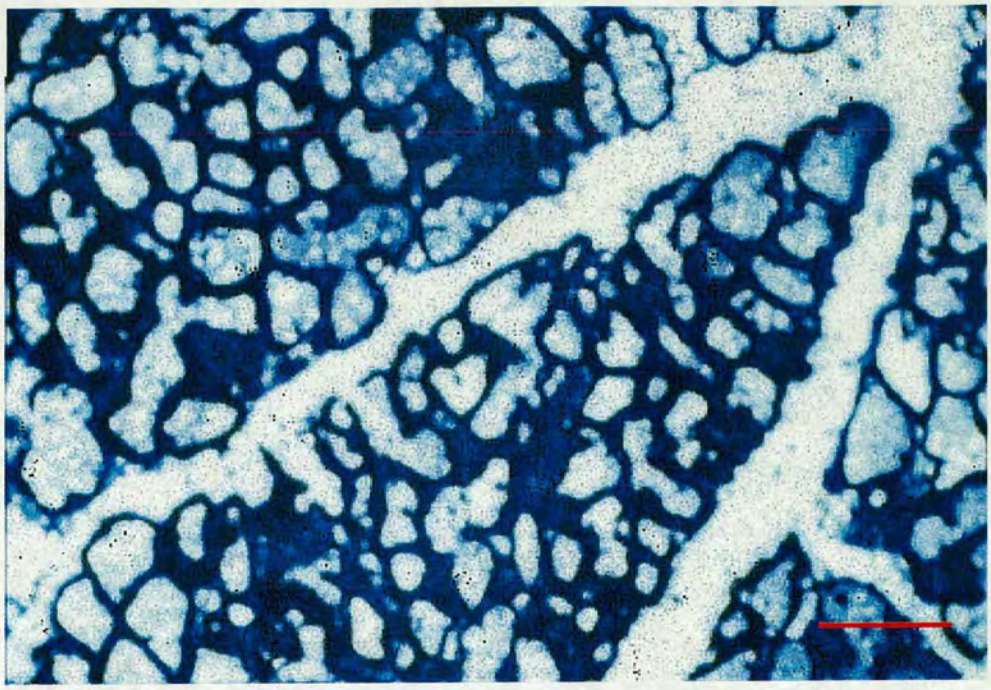


Figure 3.16. Control *in situ* hybridisation of Lox 4. (A,B,C) probed with sense probe. Photographs were taken after 2 weeks exposures using a x10 objective. (A) is a light field image. (B) is a dark field image. (C) is a composite image of both light and dark field with the reflective silver grains being represented with a yellow paint. Bar represents 300μm.

A.



B.

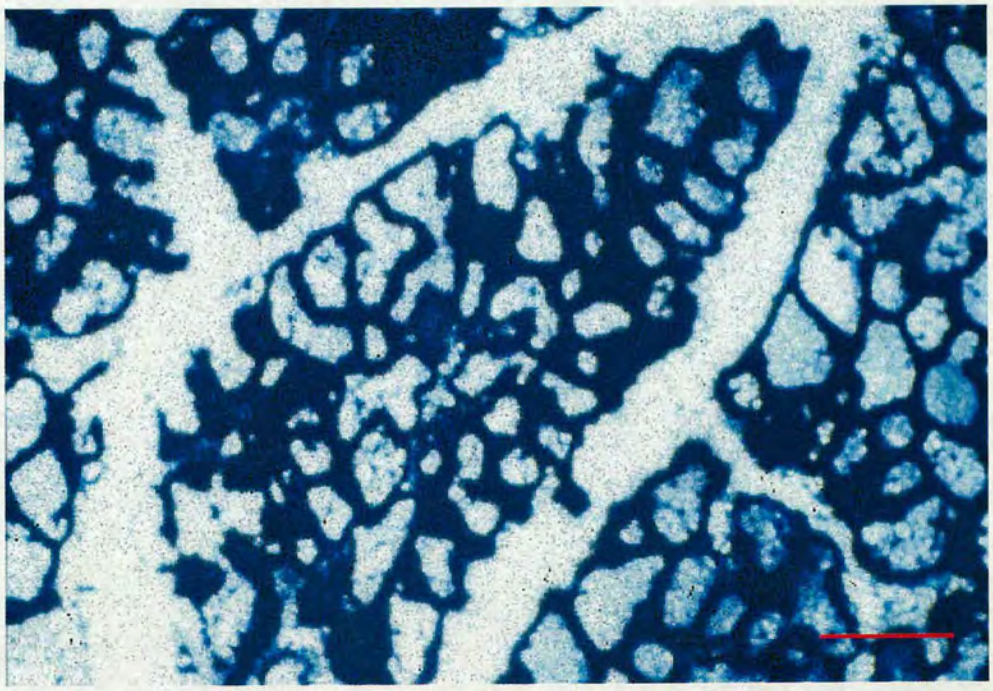


Figure 3.17. *In situ* hybridisation analysis of mRNA expression patterns in line BLG-lox 5.

(A) β-casein, (B) BLG. Photographs were taken after 2 week exposures using a x10 objective. Bar represents 300μm.

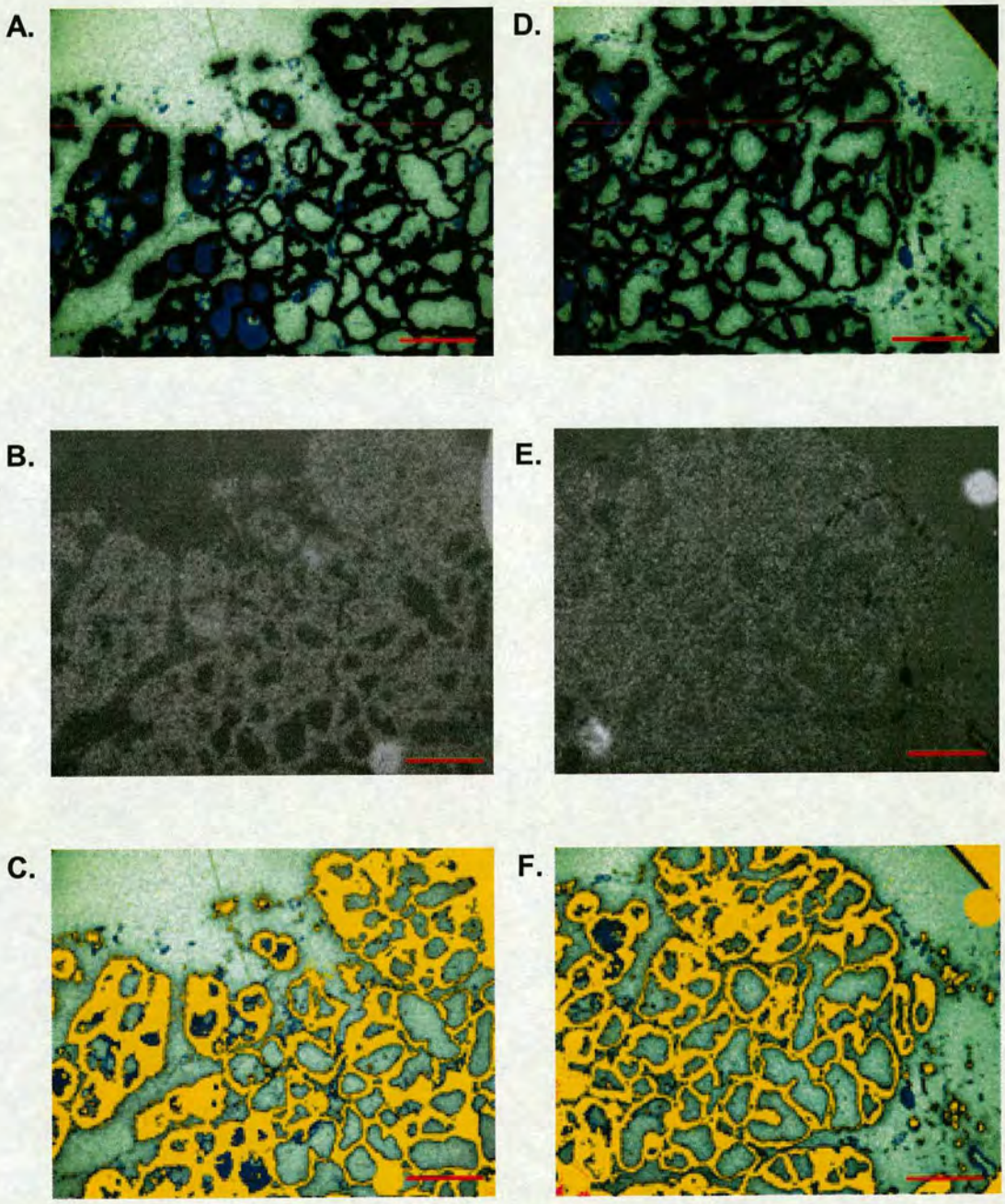


Figure 3.18. Dark field imaging of *in situ* hybridisation of Lox 5. (A,B,C) probed with β-casein. (D,E,F) probed with BLG. Photographs were taken after 2 weeks exposures using a x10 objective. (A) and (D) are light field images. (B) and (D) are dark field images. (C) and (F) are composite images of both light and dark field with the reflective silver grains being represented with a yellow paint. Bar represents 300μm.

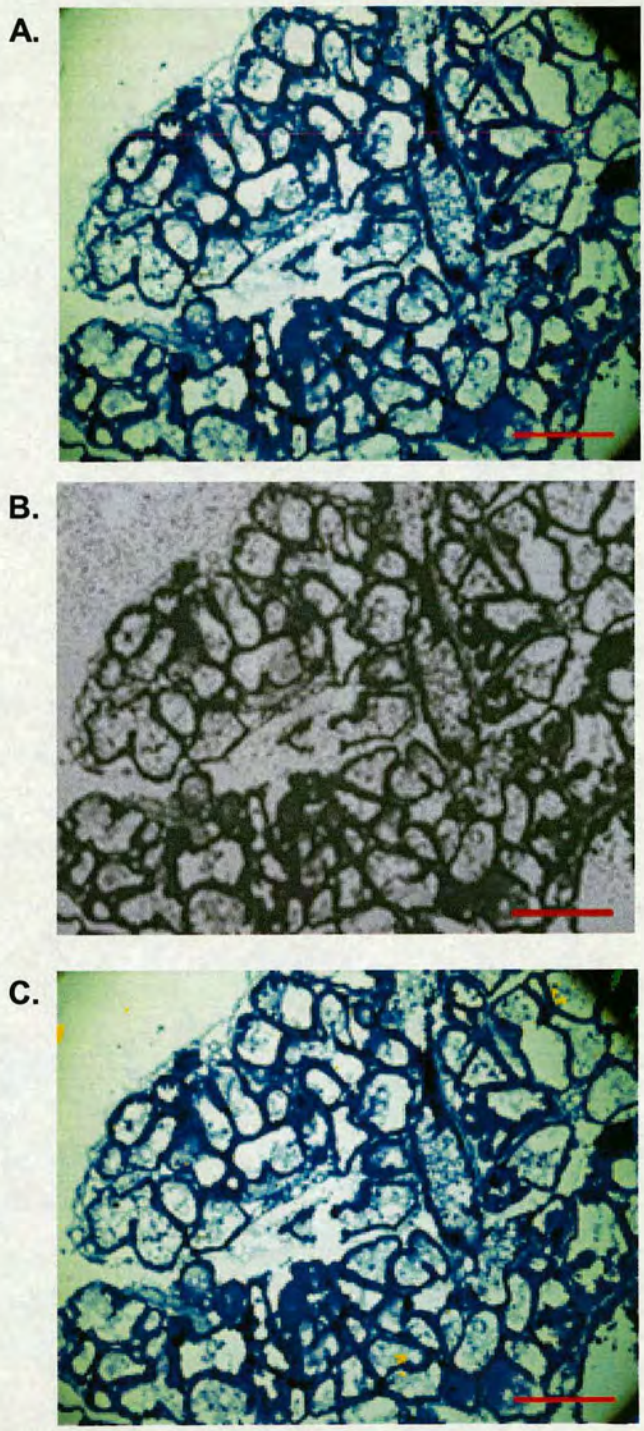
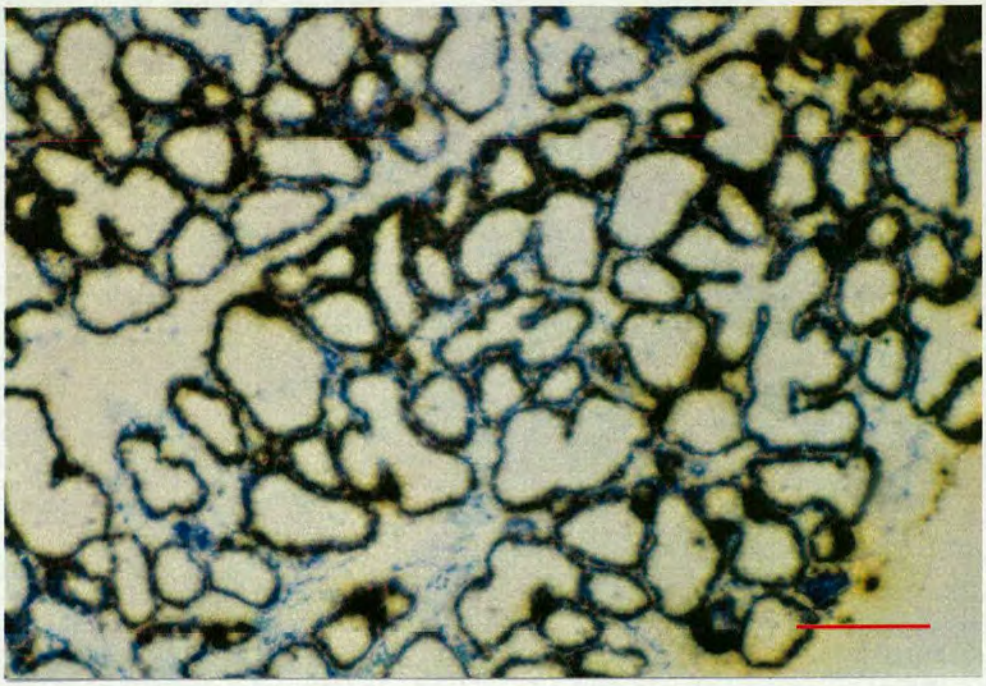


Figure 3.19. Control *in situ* hybridisation of Lox 5. (A,B,C) probed with sense probe. Photographs were taken after 2 weeks exposures using a x10 objective. (A) is a light field image. (B) is a dark field image. (C) is a composite image of both light and dark field with the reflective silver grains being represented with a yellow paint. Bar represents 300μm.

A.



B.

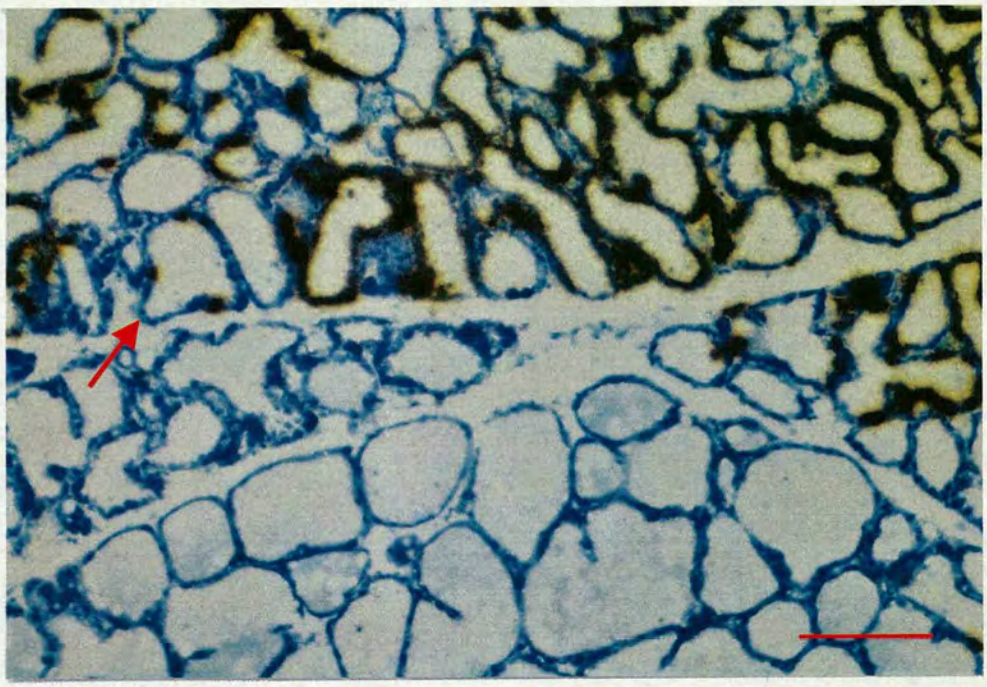


Figure 3.20. *In situ* hybridisation analysis of mRNA expression patterns in line BLG-lox 9.

(A) β-casein, (B) BLG. Photographs were taken after 2 week exposures using a x10 objective. Bar represents 300μm. Red arrow indicates an alveolus with partial expression.

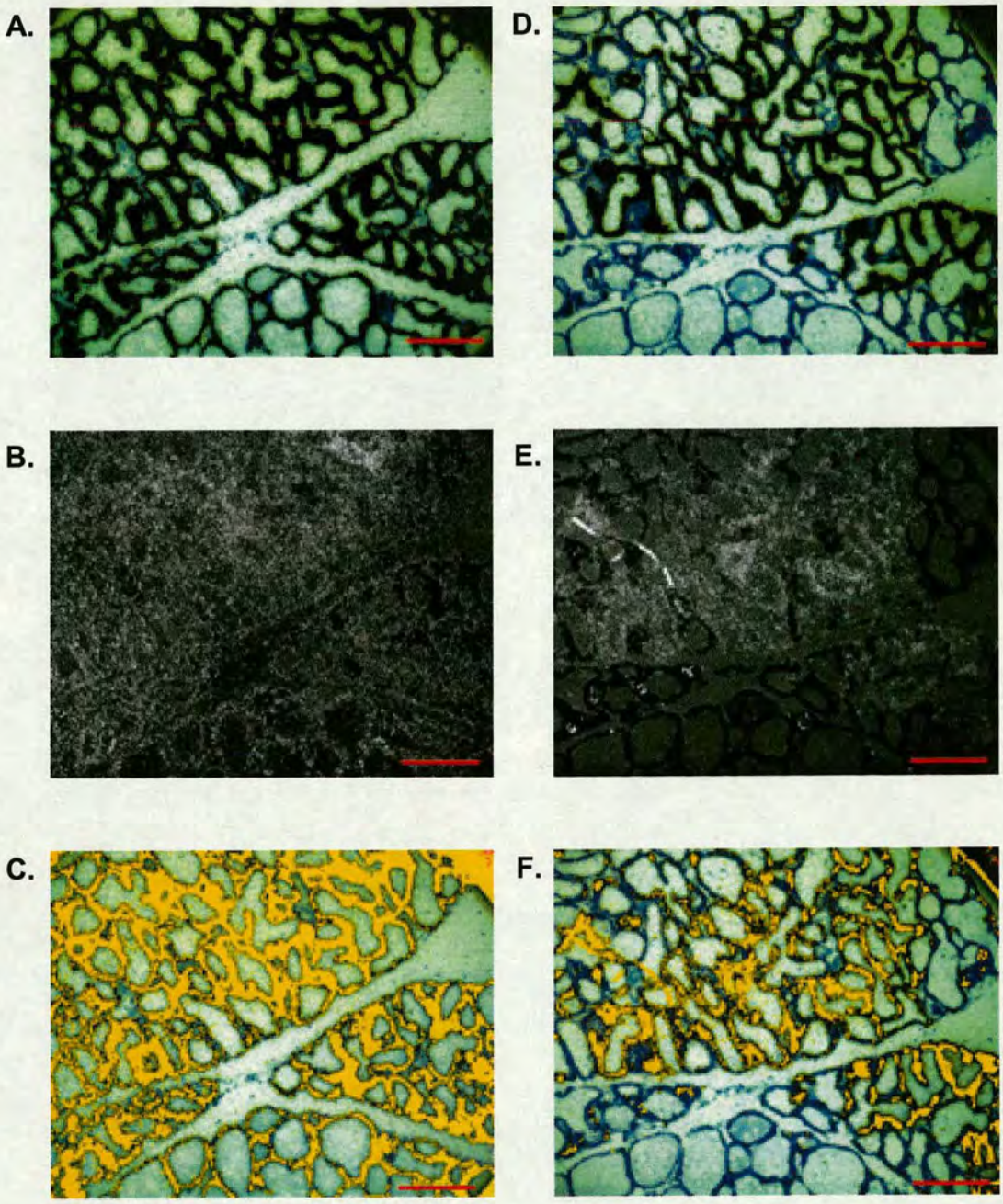


Figure 3.21. Dark field imaging of *in situ* hybridisation of Lox 9. (A,B,C) probed with β-casein. (D,E,F) probed with BLG. Photographs were taken after 2 weeks exposures using a x10 objective. (A) and (D) are light field images. (B) and (D) are dark field images. (C) and (F) are composite images of both light and dark field with the silver grains being represented with a yellow paint. Bar represents 300μm.

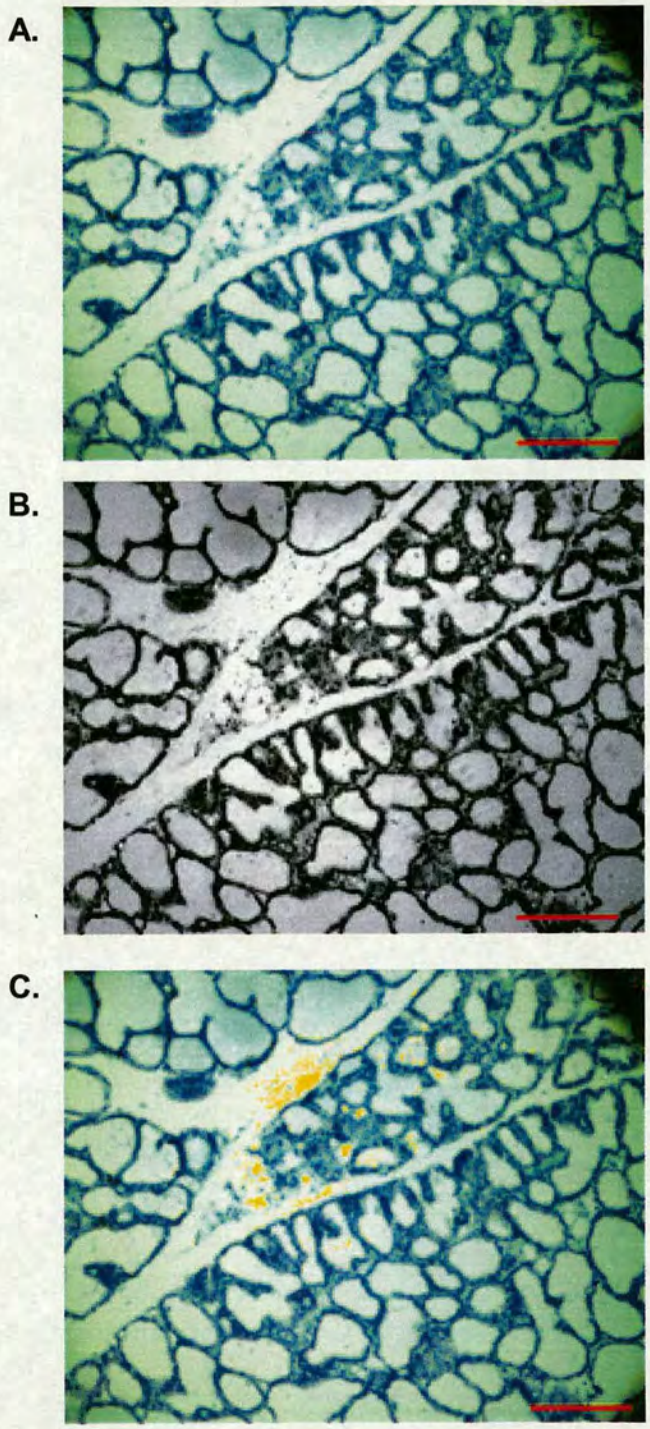
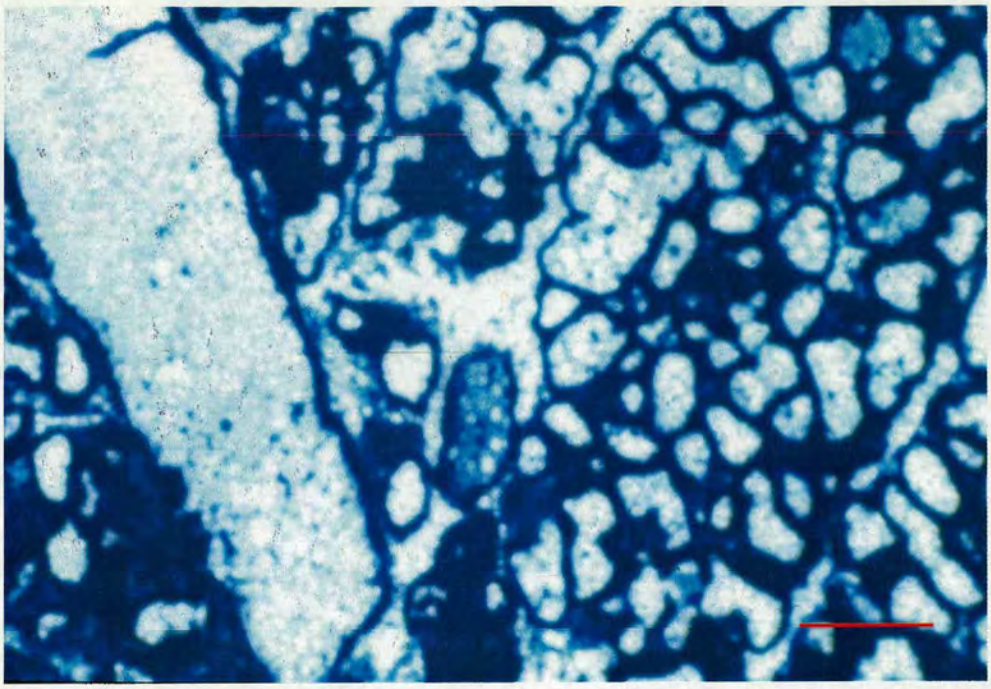


Figure 3.22. Control *in situ* hybridisation of Lox 9. (A,B,C) probed with sense probe. Photographs were taken after 2 weeks exposures using a x10 objective. (A) is a light field image. (B) is a dark field image. (C) is a composite image of both light and dark field with the silver grains being represented with a yellow paint. Bar represents 300μm.

A.



B.

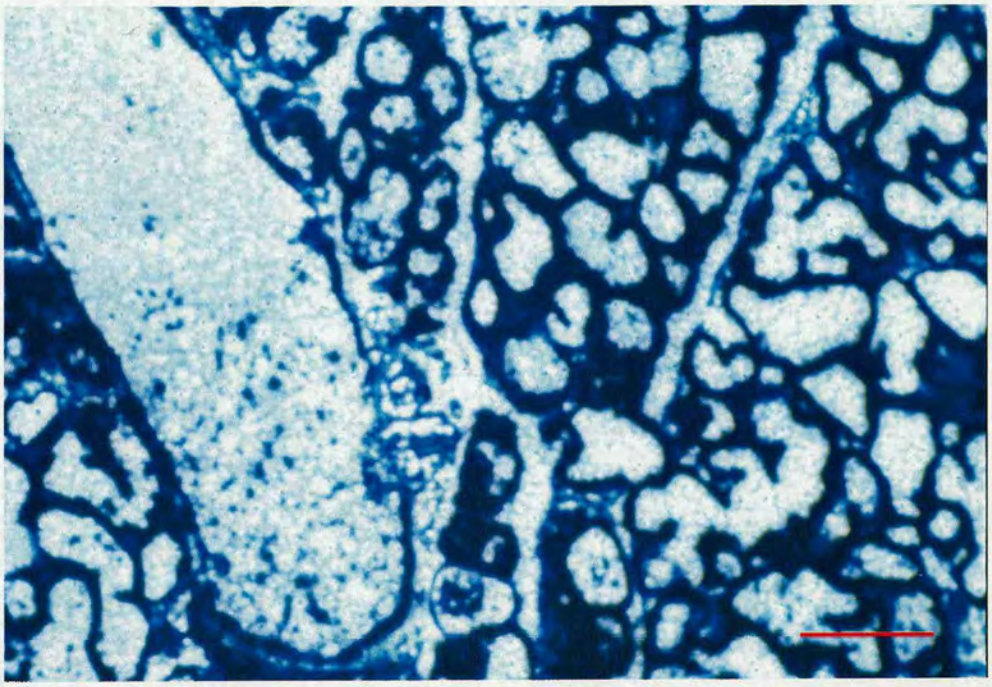


Figure 3.23. *In situ* hybridisation analysis of mRNA expression patterns in line BLG-lox 50.

(A) β-casein, (B) BLG. Photographs were taken after 2 week exposures using a x10 objective. Bar represents 300μm.

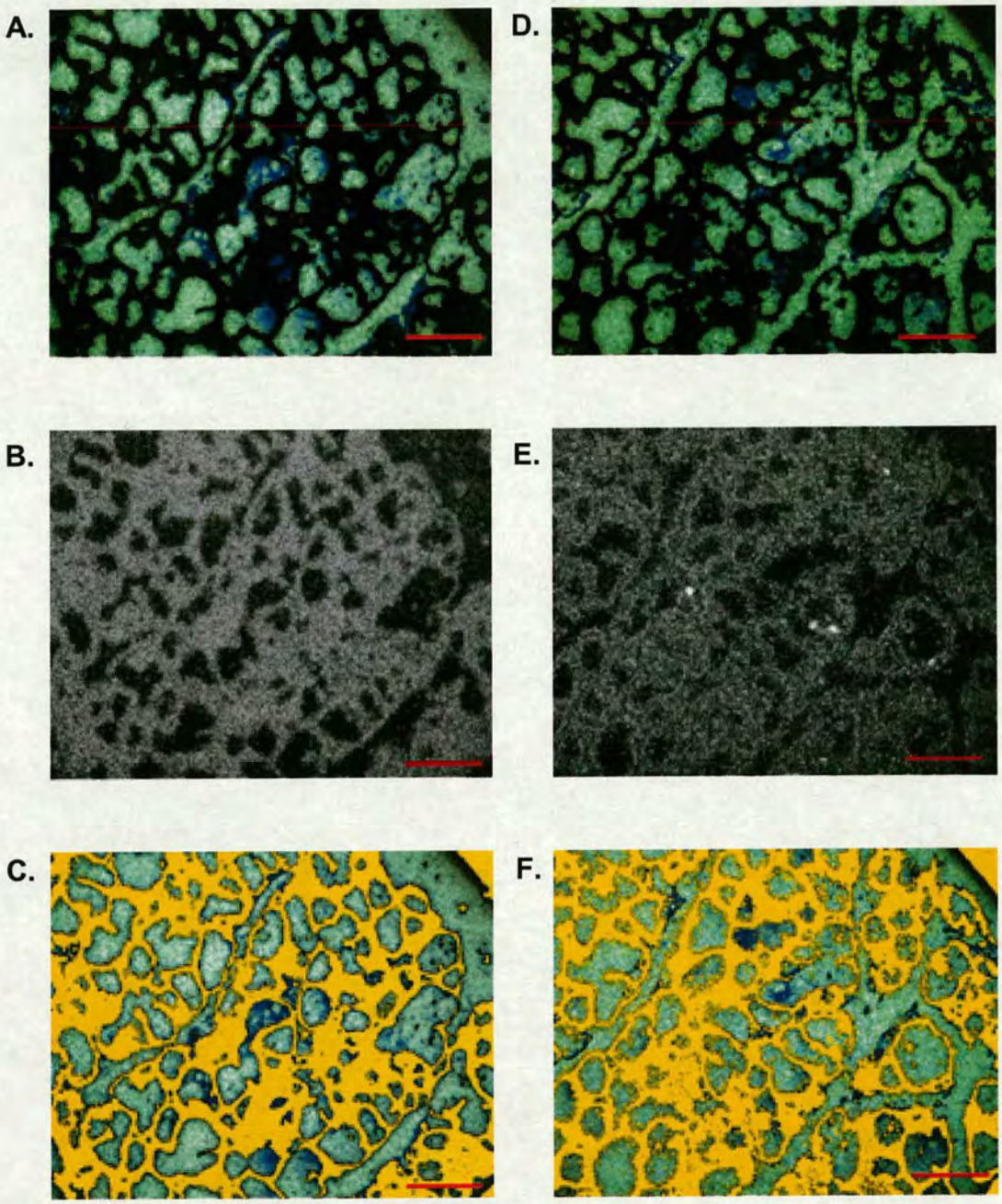


Figure 3.24. Dark field imaging of *in situ* hybridisation of Lox 50. (A,B,C) probed with β-casein. (D,E,F) probed with BLG. Photographs were taken after 2 weeks exposures using a x10 objective. (A) and (D) are light field images. (B) and (D) are dark field images. (C) and (F) are composite images of both light and dark field with the silver grains being represented with a yellow paint. Bar represents 300μm.

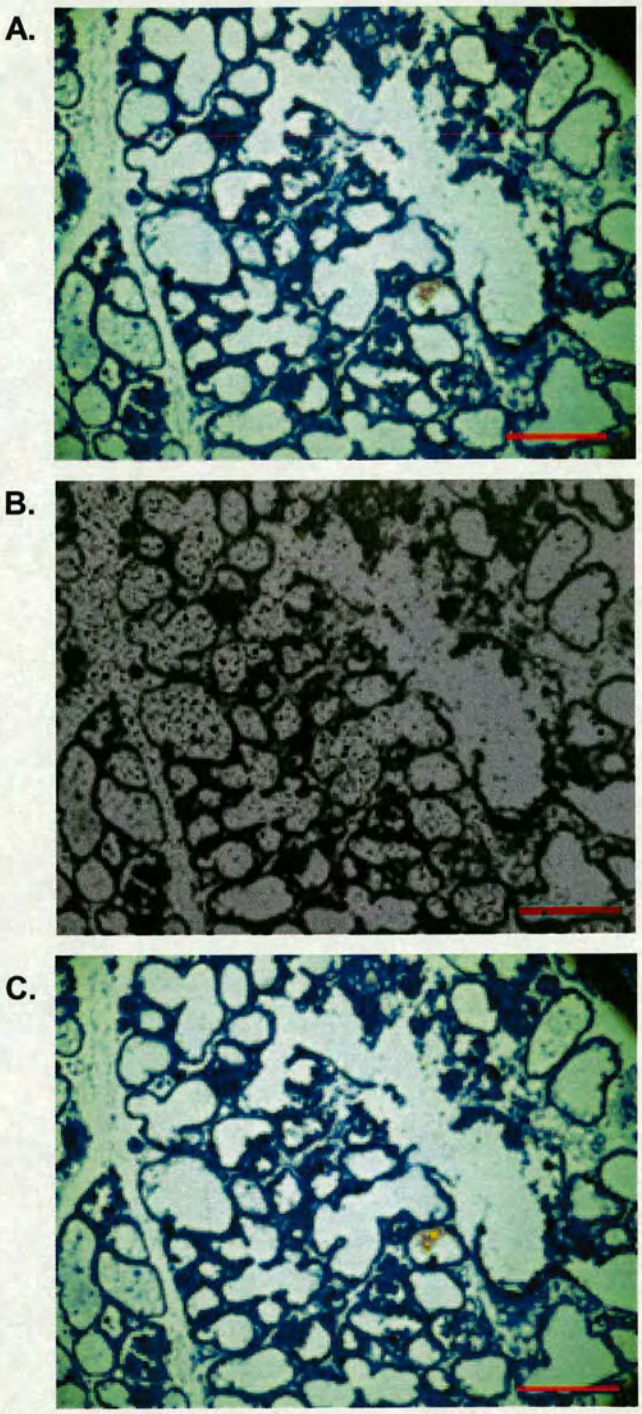
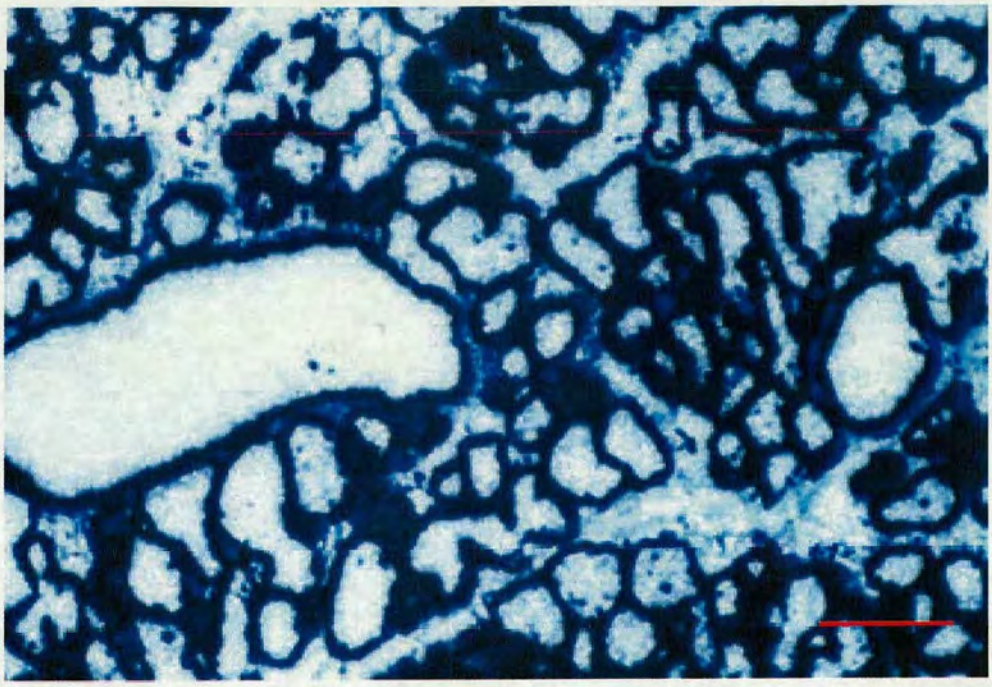


Figure 3.25. Control *in situ* hybridisation of Lox 50. (A,B,C) probed with sense probe. Photographs were taken after 2 weeks exposures using a x10 objective. (A) is a light field image. (B) is a dark field image. (C) is a composite image of both light and dark field with the silver grains being represented with a yellow paint. Bar represents 300μm.

A.



B.

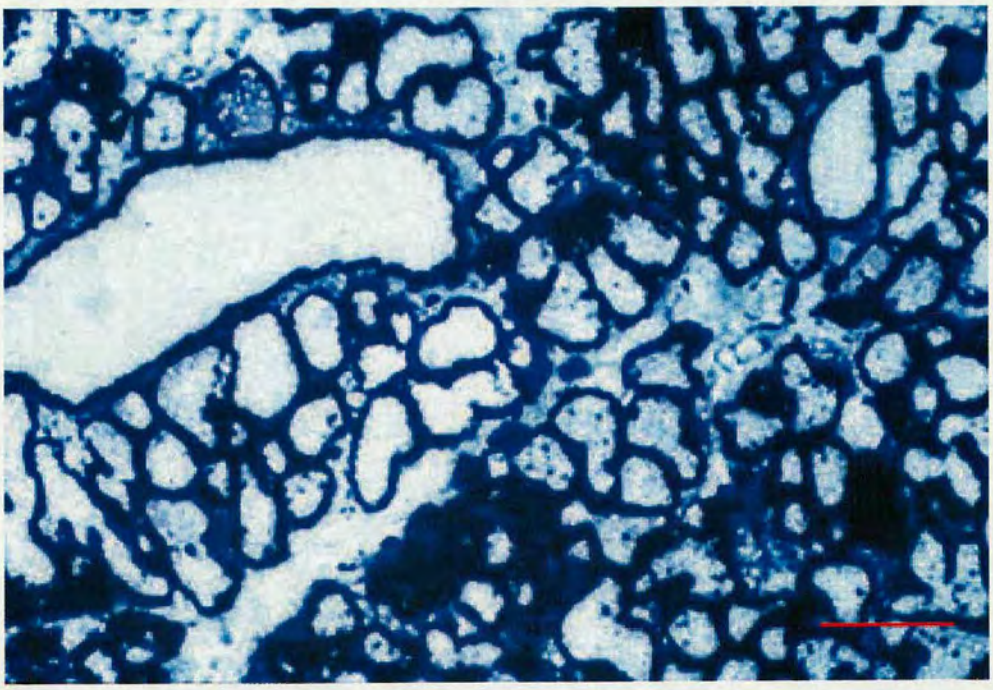


Figure 3.26. *In situ* hybridisation of mRNA expression patterns in line BLG-loxP 96.

(A) β-casein, (B) BLG. Photographs were taken after 2 week exposures using a x10 objective. Bar represents 300μm.

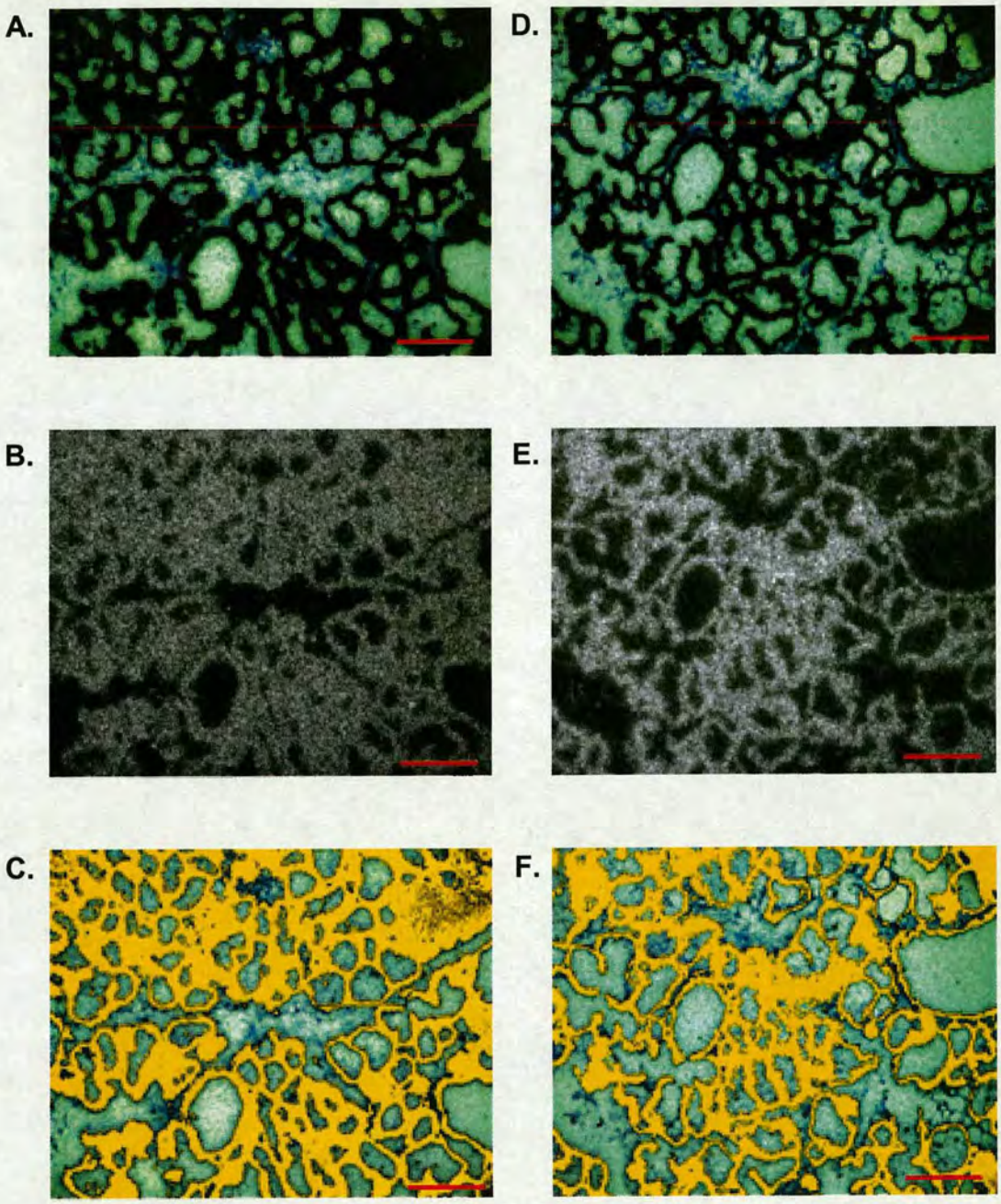


Figure 3.27. Dark field imaging of *in situ* hybridisation of Lox 96. (A,B,C) probed with β-casein. (D,E,F) probed with BLG. Photographs were taken after 2 weeks exposures using a x10 objective. (A) and (D) are light field images. (B) and (D) are dark field images. (C) and (F) are composite images of both light and dark field with the silver grains being represented with a yellow paint. Bar represents 300μm.

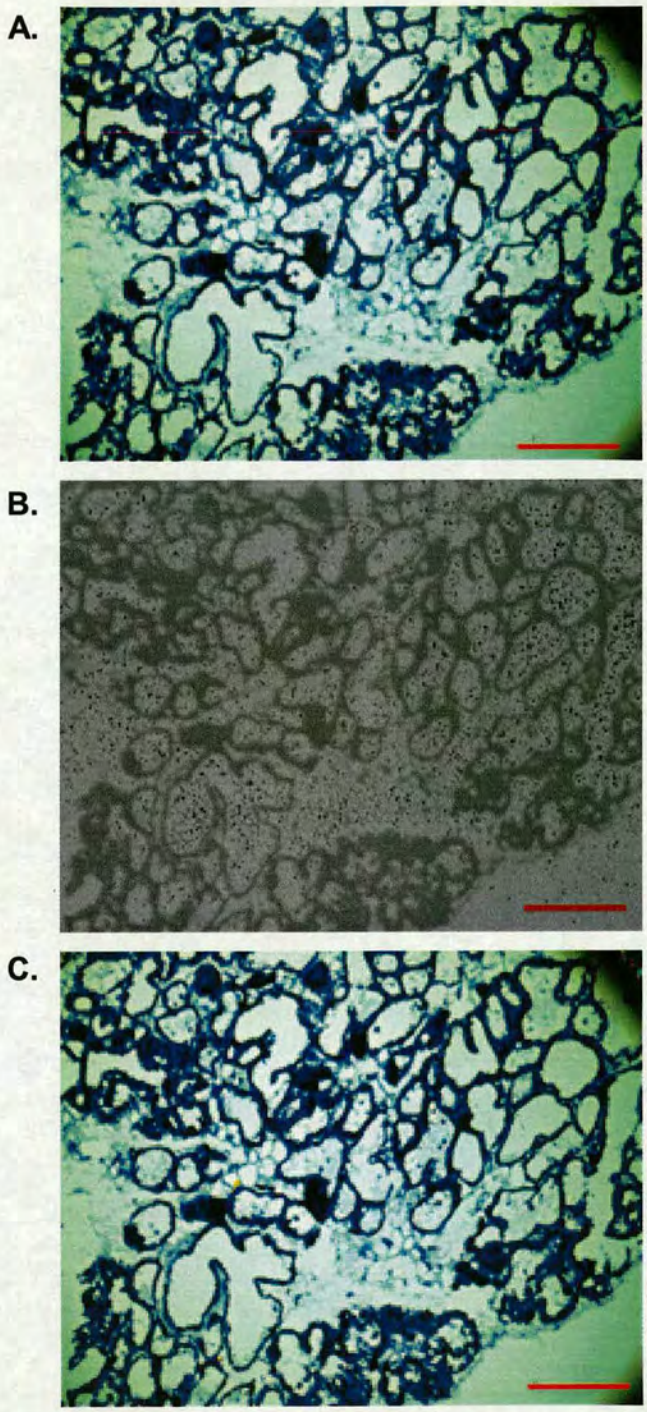


Figure 3.28. Control *in situ* hybridisation of Lox 96. (A,B,C) probed with sense probe. Photographs were taken after 2 weeks exposures using a x10 objective. (A) is a light field image. (B) is a dark field image. (C) is a composite image of both light and dark field with the silver grains being represented with a yellow paint. Bar represents 300μm.

Steps were taken to estimate the degree of heterogeneity of expression detected by *in situ* hybridisation for lox 4, 5 and 9 animals. Alveoli were scored according to the area expressing BLG RNA; negative when no cells in the alveolus expressed, -/+ when there was partial expression from an alveolus (an example is indicated in figure 3.20 by a red arrow), and positive when all the cells in an alveolus expressed. Two sections per animal were counted at x20 magnification using ten fields of view. This was to try and avoid bias in the areas analysed, as the ten fields generally covered the whole of the section. Other factors including orientation on the slide, parentage and litter number were assumed to have no effect on the outcome.

Figures giving the percentage of alveoli in each group are given in table 3.2. The counts show that in the lox 4 and 9 lines, alveoli fall into all three categories. In these two lines there is variation between individuals within the same transgenic line, in the number of cells expressing. In lox 5 all the alveoli counted showed uniform expression of BLG.

The variation within lox 4 and 9 animals is shown in figure 3.29 which is a dotplot of the percentage of alveoli that are positive (partial and full expression) in each individual mouse. Lox 5 is in constrast with this, as all animals from this line are found to express BLG mRNA in all the alveoli examined and counted.

The counts were analysed using GenStat software to produce Table 3.3 combining the values of alveoli counted as partly positive and those fully positive. Standard deivations (s.d.) and coefficient of variance are both measures of the amount of variability within a group i.e. the differences between individuals from the same transgenic line.

Animal	*Negative alveoli*	*Partly positive alveoli*	*Positive alveoli*
Lox 4.1	40	36	24
Lox 4.3	45	27	29
Lox4a5.20.4	43	37	20
Lox4a5.20.5	56	26	18
Lox4a5.20.7	46	23	31
Lox4a5.20.8	47	22	31
Lox5.7	0	0	100
Lox5.10	0	0	100
GB2.9	0	0	100
Lox9.2	23	24	53
Lox9.4	44	27	29
GB11.5	55	21	24

Table 3.2. BLG expression values by *in situ* hybridisation of lox lines. Ten fields per section and two sections for each animal were counted to derive a percentage mean for the three categories.

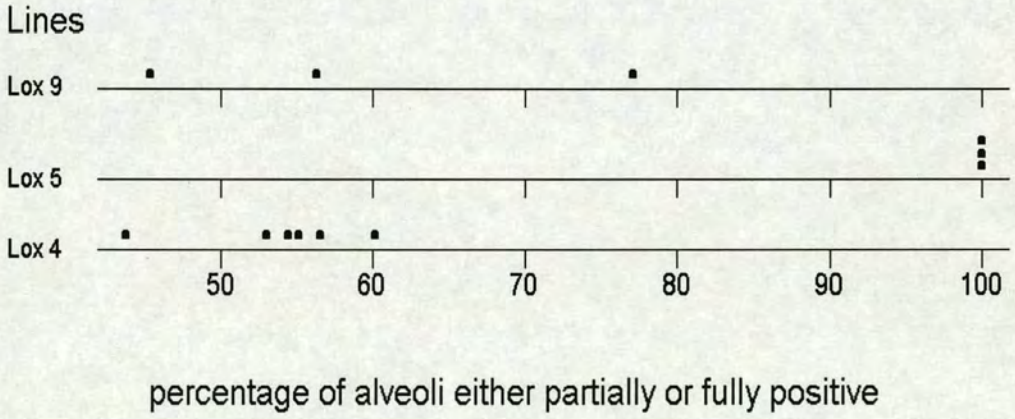


Figure 3.29. Dotplot of combined positive alveoli for lox animals. Each dot represents an animal. The percentage of alveoli that were counted partial or fully positive were combined to give an overall view of expression in the mammary gland.

Line	*Range of percentages of positive alveoli*	*n*	*Mean*	*s.d*	*c.v*
Lox 4	44-60	6	54	5.5	10.3
Lox 5	100-100	3	100	0	0
Lox 9	45-77	3	59	16.2	27.1

Table 3.3. Statistical analysis of lox *in situ* counts. Range is given as a percentage of the total alveoli which had some expression, *n* = number of animals per line, *s.d.* standard deviation, *c.v.* = 100 x s.d. / mean.

As expected from the *in situ* figure 3.17 there is no variation among the individuals of the stably expressed lox 5 line. However, both lox 4 and 9 show variation between individuals at a cellular level in the number of cells which express the BLG transgene, and this is reflected in a high coefficient of variance. Only three animals were counted for the lox 5 and 9 lines. However, despite the small sample size there are clear differences between the lines. More animals (n=6) were analysed for the lox 4 line as this was used more extensively thoroughout the other chapters, and therefore enabled more samples to be analysed.

3.2.7.2 Northern blot analysis

Northern blotting was used to confirm the presence of BLG RNA found by *in situ* hybridisation. Total RNA isolated from mid-lactation mammary tissue from the five lines all showed BLG RNA (figure 3.30). As there were only a small number of samples per line statistical analysis would not have been informative. The Northern blot does however indicate that a BLG RNA of the expected size is produced in the lox transgenic animals.

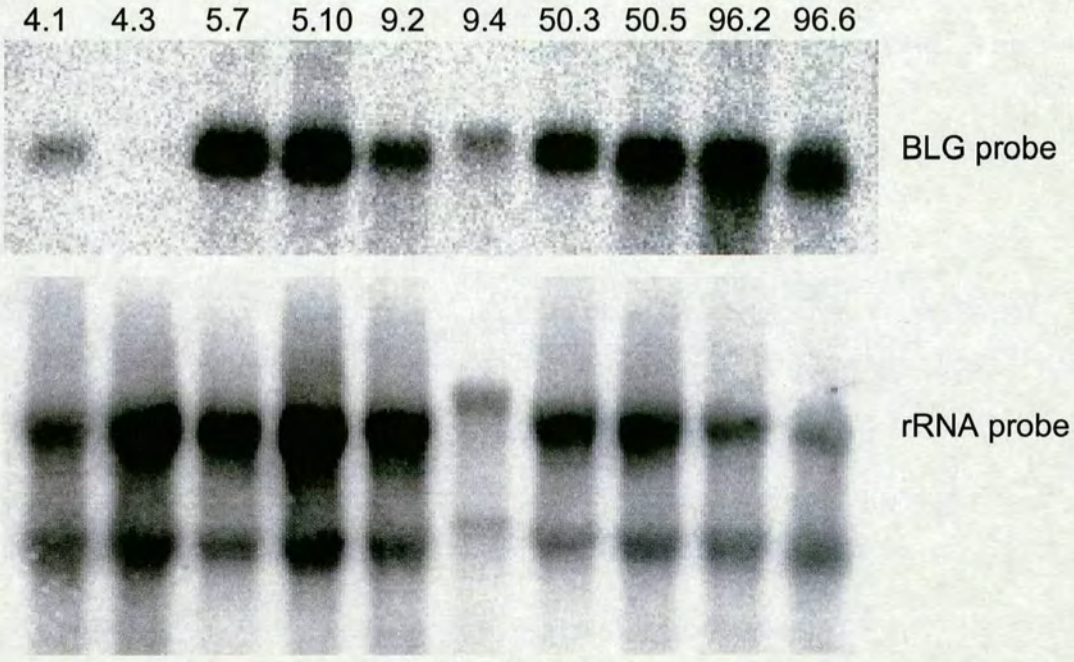


Figure 3.30. Northern blot of lox mammary RNA. Total RNA was electrophoresed on a denaturing gel, blotted to Zetaprobe, and then sequentially probed with a BLG cDNA PCR product and a 28S 12kb fragment to control for loading differences.

3.2.7.3 Milk Protein analysis

SDS-PAGE analysis was used to determine if BLG protein was present in transgenic milk samples from the five lox lines (figure 3.31). Mouse milk samples were prepared as described (see 2.7.1) and were loaded at 1/200 dilutions along side ovine BLG standards (a gift from Dr. Maggie McClenagahan; Roslin Institute). Gels were photographed using Quantity One software (BioRad). From SDS-PAGE analysis it could be seen that all of the transgenic animals expressed BLG protein. As lox 4, 5 and 9 were to be used in the reduction experiments, a larger number of individual milk samples were each run on triplicate gels to give accurate measurements. Standard deviations and coefficients of variation were calculated in table 3.4.

Line	*Range of BLG protein in line*	*n*	*Mean*	*s.d.*	*c.v*
Lox 4	0.7-1.5	8	1.2	0.3	24
Lox 5	8.3-10.6	6	9.5	0.8	8
Lox 9	3.0-8.1	7	5.2	1.9	37

Table 3.4. BLG protein expression levels in lox 4, 5 and 9. BLG protein values are mg/ml, *n* = number of animals analysed per line, *s.d.* standard deviation, *c.v.* = 100 x s.d. / mean.

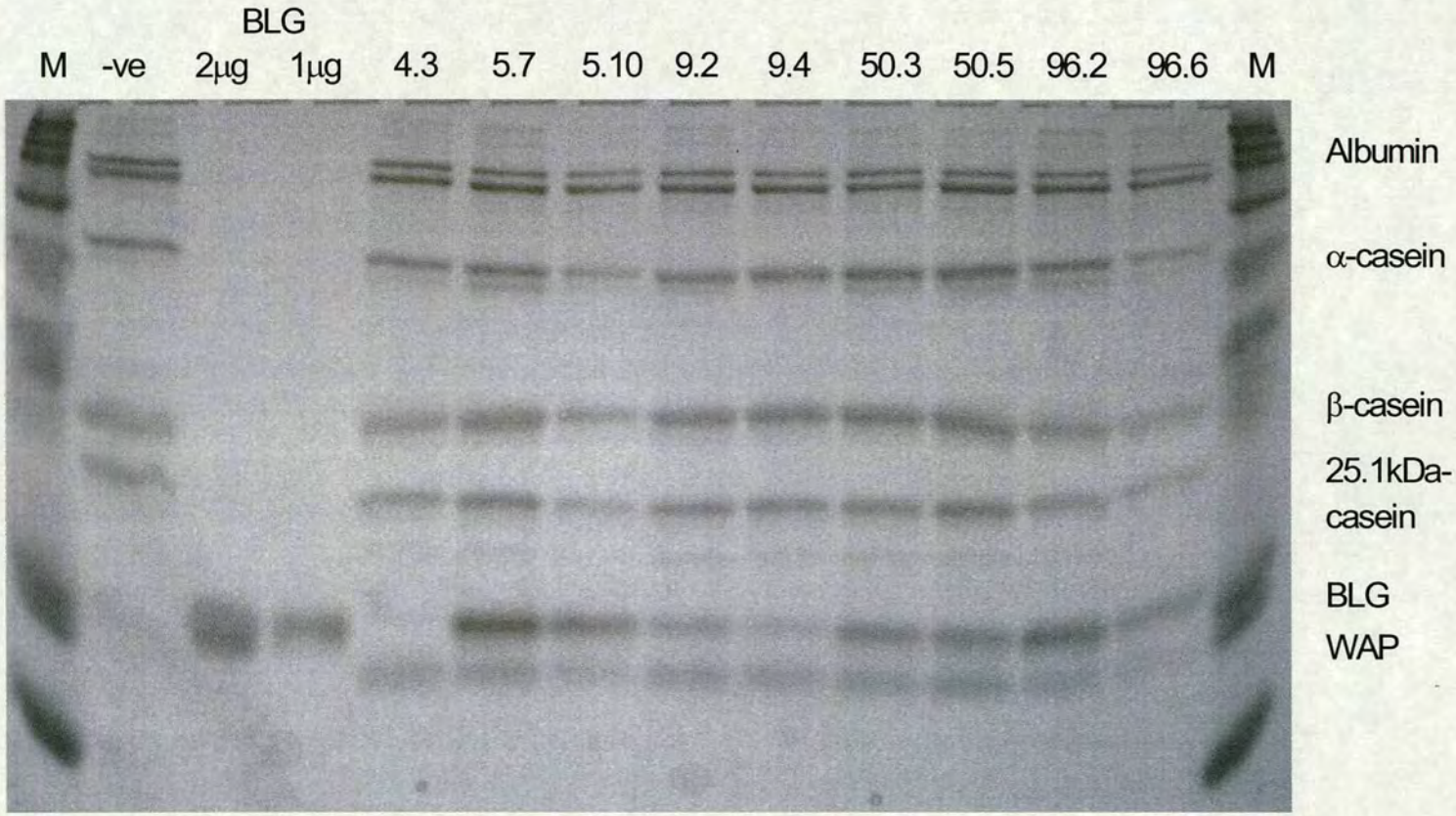


Figure 3.31. SDS-PAGE analysis of lox milk proteins. Mouse milk samples were electrophoresed in 18% polyacrylamide gels under reducing conditions and stained with Coomassie Blue. M indicates Rainbow marker (Amersham). -ve indicates a 1/200 dilution of defatted control non-transgenic mouse milk sample. 1/200 dilutions of defatted milk from transgenic animals from each of the five lox lines was loaded along side 1μg and 2 μg of BLG.

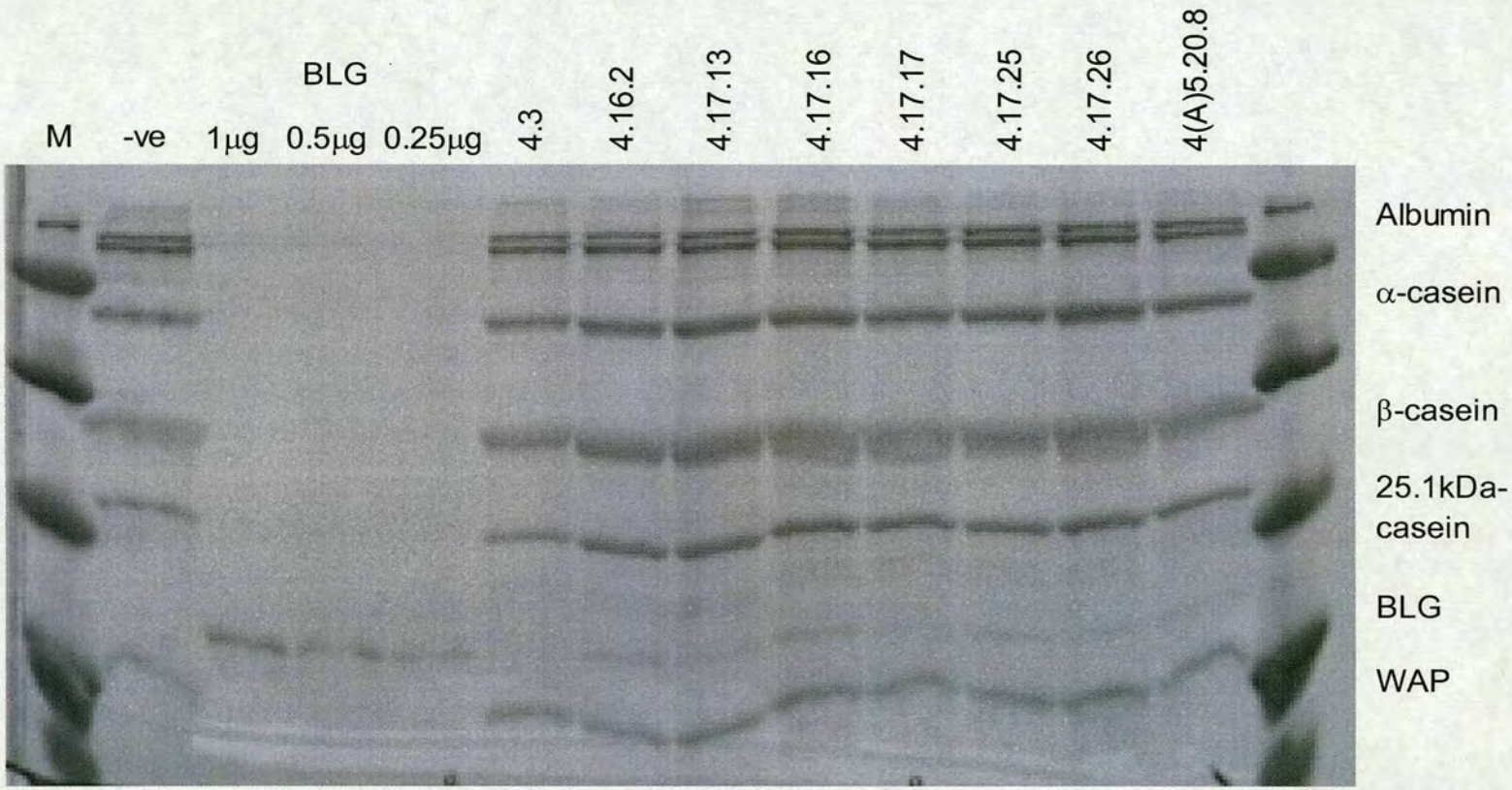


Figure 3.32. SDS-PAGE analysis of lox 4 milk proteins. Mouse milk samples were electrophoresed in 18% polyacrylamide gels under reducing conditions and stained with Coomassie Blue. M indicates Rainbow marker (Amersham). -ve indicates a 1/150 dilution of defatted control non-transgenic mouse milk sample. 1/150 dilutions of defatted milk from transgenic lox 4 animals were loaded along side 1μg, 0.5μg and 0.25 μg of BLG.

These further investigations of lox 4 indicated that the BLG protein level did indeed vary between animals from that line (figure 3.32). The values from lox 4 showed a range of 0.70– 1.5 mg/ml. The levels between individuals in the lox 9 line also showed variation in the level of BLG protein, with over a 2 fold difference between samples. With more lox 5 samples available it was possible to show the stable nature of the BLG concentration in this line, as illustrated by the lower *c.v* from this line. All protein analysis confirmed the nature of the lox lines as established by the *in situ* data.

Previous studies on BLG transgene variegation also analysed the BLG protein concentration within a line. One variegating line showed higher levels of protein expression than any of the lines analysed here (23.7mg/ml in BLG 45 line; Dobie *et al* 1996). Higher levels of variation between sibs was found previously: BLG line 7 had a *c.v.* of 52 (Dobie *et al* 1996), compared to lox 9 with a c.v. of 37 (table 3.4).

The stable lox 5 line in this study, showed a lower amount of variation (*c.v.* of 8) compared with the uniformly expressed line 14 (*c.v.* of 17) in the previous work (Dobie *et al* 1996). This may be due to the differences in the methods to quantify the visible bands from SDS-PAGE. In the previous study bands were quantified by scanning densitometry, print outs were used to cut out the peak representing the BLG band and these were then weighed with a fine balance; in this study gels were scanned using a BioRad MultiImager and Quantity One software.

Western blotting of SDS-PAGE gels was carried out to confirm the identity of the novel band found in transgenic milk samples. Immunoblotting of the lox milk proteins confirmed that the novel band

found in the transgenic milk contained an antigen that reacted to BLG antibody (figure 3.33). In the lox 5 samples due to the higher level of protein the bands are diffuse due to overloading of the gel.

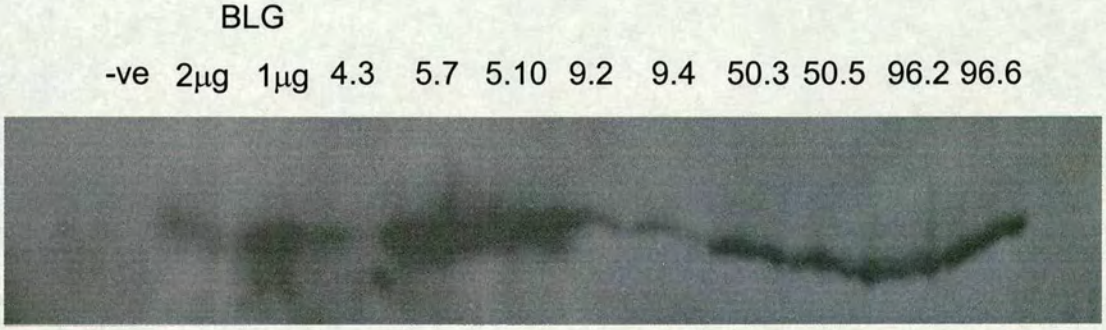


Figure 3.33. Western blot of lox milk. 1/200 dilutions of milk protein samples were run on a SDS-PAGE gel, electroblotted to nitrocellulose and probed with rabbit anti – BLG and then HRP conjugated anti rabbit IgG.

3.3 CONCLUSIONS

A number of transgenic mouse lines were generated with a BLG transgene containing a single loxP inserted in the 3′ flanking sequence. Southern blotting was carried out to determine the copy number and the structure of the transgenes integrated in each of the lox lines. All had more than 10 transgene copies integrated. In lox 4, 5 and 50 the array was determined to be a primarily simple head to tail arrangement. Lox 9 was more complex, while lox 96 had an extremely complex array which may have been generated by two separate but closely linked integrations. These transgenic arrays were stable; there was no change in copy number/structure between generations or within litters.

In situ hybridisation analysis showed that out of the five high copy number lines generated, two (lox 4 and 9) showed variegated BLG expression, while the other three showed a uniform pattern of BLG expression.

Data from the SDS-PAGE protein analysis correlated with the estimation of *in situ* expression, indicating that lines with the highest numbers of cells expressing BLG mRNA also showed the highest levels of BLG protein in milk samples.

The transgenic mice studied for this project were of a C57BL/6 x CBA hybrid genetic background. Background modifiers have previously been identified that interfere with transgene expression (Allen *et al* 1990). It is possible that the different levels of expression observed among individuals

within a lox line may be due to segregation of alleles of a modifier gene, which interferes with transgene expression. Previously variegating BLG transgenes on a mixed genetic background, when backcrossed to 97% CBA or C57BL/6 backgrounds showed no difference in the variability of transgene expression, indicating modifiers were not responsible for the varigation within a BLG transgenic line (Dobie *et al* 1996). However, this thesis did not attempt to address whether genetic modifiers were influencing the BLG-loxP transgene expression patterns. Without conducting backcrossing experiments on the lox lines, modifier genes cannot formally be ruled out as a possible reason for the variegation within the lox lines.

CHAPTER FOUR

TARGETED REDUCTION OF TRANSGENIC COPY NUMBER IN THE MAMMARY GLAND

4.1 INTRODUCTION

Having established five transgenic lox lines and analysed their expression patterns as discussed in Chapter 3, this chapter focuses on the reduction of transgene copy number specifically in the mammary gland and investigates the effect on the expression profile of the transgene.

Cre recombinase is a site specific recombinase which causes the deletion of DNA between two loxP sites in the same orientation (Sauer and Henderson 1989). It has now been extensively used in transgenic animals for a variety of purposes as discussed in Chapter 1. Cre recombinase has been expressed by using a number of tissue-specific promoter to drive expression to the tissue of interest. In this case BLG-Cre (Cre 74) transgenic mice had previously been generated in this lab and had been tested against a lox reporter construct. Cre expression was shown to be specific to the secretory epithelial cells of the mammary gland and was expressed in a uniform manner when investigated using *in situ* hybridisation (Selbert *et al* 1998).

Due to the nature of the integrations of the lox lines as head to tail repeats, when Cre recombinase is introduced a deletion event should occur between the two loxP sites of the adjacent transgenes producing a reduction in copy number. To achieve this mammary-specific reduction in copy number of the lox lines, animals were crossed to a BLG-Cre line (Cre 74). Three lines were used for this further study; lox 4 and 9 which had variegated expression and lox 5 which had uniform expression.

Double transgenic animals were analysed for changes in copy number by Southern blotting, while expression analysis was carried out as before using *in situ* hybridisation, Northerns and milk protein analysis. The results are discussed below.

4.2 RESULTS

4.2.1 Production of double transgenic animals

Lox animals from lox 4, 5 or 9 were bred to Cre 74 animals and the resulting offspring were tested for the presence of both the BLG-loxP transgene and the BLG-Cre transgene by Southern blotting. Within litters animals could inherit neither transgene, the lox transgene, the Cre transgene or both transgenes (examples from the BLG-lox4 crossed with BLG-Cre are shown in figure 4.1; the same patterns were seen in the other two crosses). Double transgenic females were used for mid-lactation analysis, while a number of female sibs carrying only the BLG-loxP construct were used as controls.

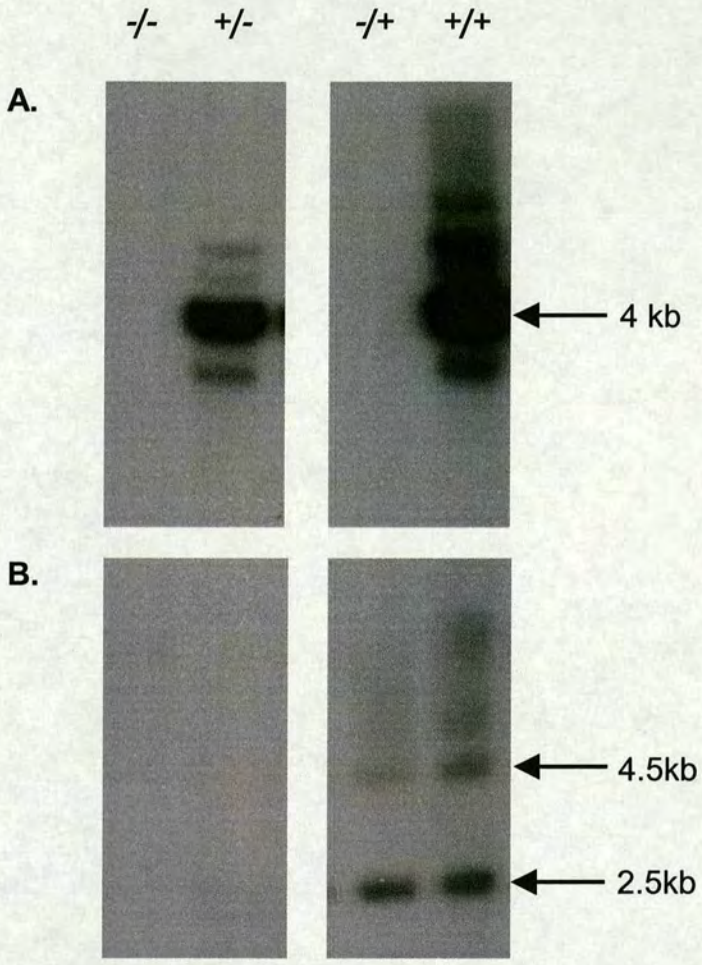


Figure 4.1. Southern blotting for presence of BLG-loxP and BLG-Cre transgenes. Progeny from crosses between BLG-loxP animals and BLG-Cre 74 animals were analysed for the presence of both transgenes. 20μg of tail DNA was digested with BamHI, the digests were electrophoresed on two 0.8% gels and blotted to Zetaprobe membrane. (A) One blot was probed with a 1.1kb BLG fragment, (B) the other blot was probed with a 1kb Cre probe. -/- indicates negative for both transgenes, +/- indicates positive for BLG-loxP only, -/+ indicates positive for BLG-Cre only, +/+ indicates positive for both transgenes.

4.2.2 Analysis of copy number reduction

The mammary gland is a complex tissue comprising more than one cell type. The mature adult mammary gland consists of secretory epithelial cells surrounded by a layer of contratile myoepithelial cells and supporting connective and adipose tissue (Mepham 1987). As the Cre recombinase used for this experiments is under the control of BLG, it is expressed in the major subset of cells in the mammary gland; the secretory epithelial cells (Selbert *et al* 1998). Therefore a proportion of the mammary gland would maintain the 'parental' transgenic array, while the majority of cells would have a 'reduced array'. This resulted in mosaicism within genomic mammary DNA which complicated the analysis of copy number reduction.

EcoRV does not have a recogition site within the BLG-loxP sequence, therefore it cannot cut within the transgenic array, but would leave the entire transgenic array as one fragment within the digested genomic DNA. This fact was used to analyse the reduction in copy number. As a high copy number array, EcoRV digestion would produce a high molecular weight fragment, which was termed the 'parental array'. If after breeding to Cre mice there had been a deletion of copies from the array it should appear as a smaller fragment, termed the 'reduced array'. This is illustrated in figure 4.2. As shown in figure 4.3 a 10kb EcoRV 'reduced array' fragment was found in lox 4 mammary gland DNA, and is referred to as the 10kb RV locus from this point. The parental array also showed a decrease in intensity. The liver tissue is used as a control for the parental state as Cre recombinase is not expressed in this tissue (Selbert *et al* 1998).

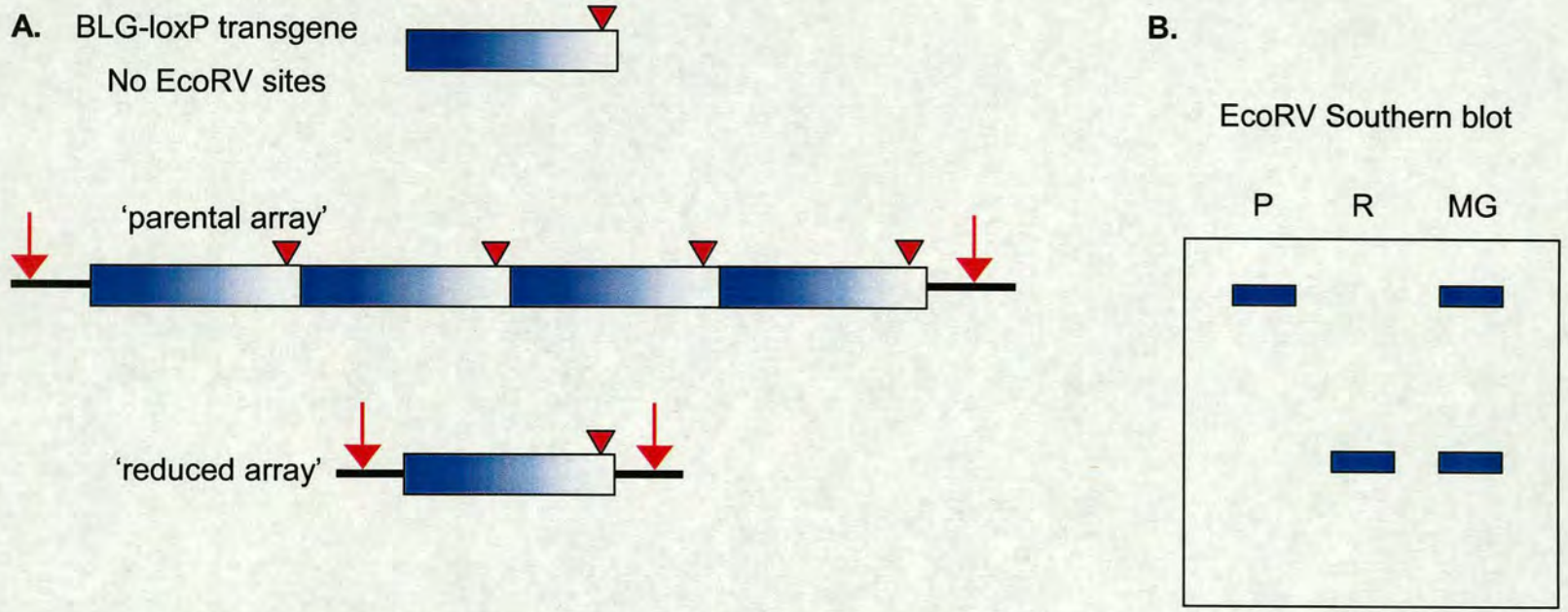


Figure 4.2. Diagram of EcoRV digestion of double transgenic DNA. (A) The BLG-loxP transgene does not contain any EcoRV sites (red arrows represent the EcoRV sites, red triangles represent loxP sites) within it. The original lox lines contain multicopy arrays. The multicopy array would not be digested by EcoRV, and would present as a high molecular weight fragment termed the 'parental' array. This would consist of all the transgene copies plus the surrounding genomic DNA until the first genomic EcoRV sites. In double transgenic animals the Cre recombinase would be expected to remove BLG-loxP transgenes to produce a 'reduced' array. This would therefore produce a smaller EcoRV digestion fragment. (B) When EcoRV digested DNA is Southern blotted, there would be an expected size difference between the two types of arrays; P represents the parental array, R represents the reduced array. Because the mammary gland consists of a mixture of cell types, some of which do not express the BLG-Cre construct, there would be mosaic pattern of both arrays represented by MG (mammary gland).

The 10kb RV locus was only ever found in lox 4 animals carrying both BLG-loxP and the BLG-Cre constructs, and furthermore, was never found in the liver of these animals, indicating that it is generated by the presence of Cre recombinase. The size of the fragment was consistent with it comprising one copy of the lox transgene plus the surrounding genomic DNA up to the genomic EcoRV sites on either side of the array (figure 4.4).

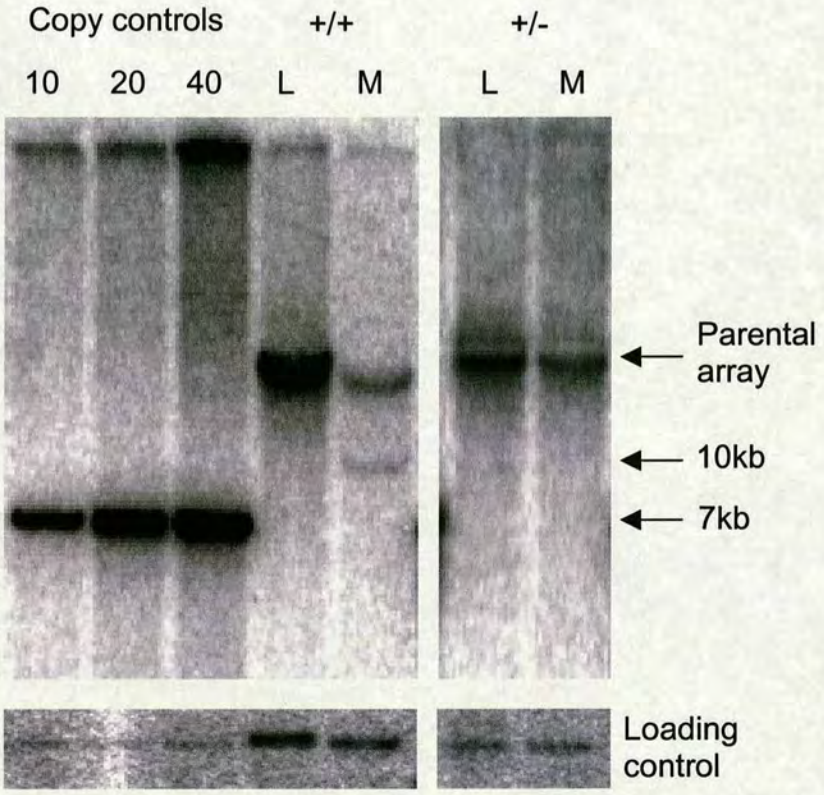


Figure 4.3. EcoRV digestion of lox 4 double transgenic DNA compared with lox 4 DNA. 10μg liver (L) or mammary (M) DNA was digested with EcoRV, electrophoresed on a 0.8% gel and blotted to Zetaprobe membrane, and then probed with a 1.1kb BLG fragment. Copy controls were dilutions of the original microinjected 7kb BLG-loxP fragment added to 10μg of non transgenic DNA. +/+ indicates an animal containing both lox 4 and Cre transgenes, +/- indicates an animal containing only the lox 4 transgene. The blot was probed with CFTR probe to control for loading differences.

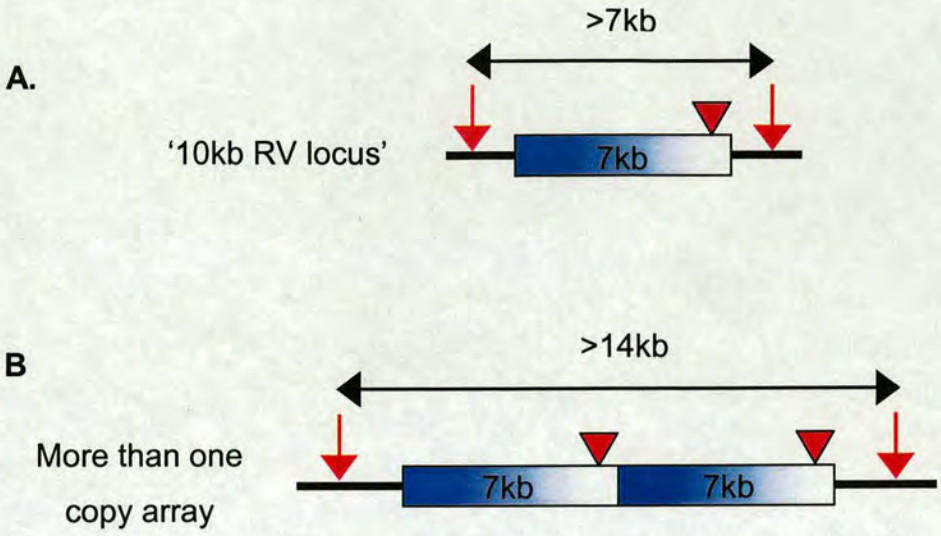


Figure 4.4. Diagram of 'reduced array'. (A) The 10kb RV locus is consistent with one copy of the lox transgene and the genomic DNA surrounding it until the first EcoRV sites (red arrows represent EcoRV sites, red triangle represent a loxP site). (B) If there were more than one transgene copy the 'reduced array' would be expected to be greater than 14kb.

The ratio of the EcoRV fragments gives an indication of the percentage of cells that have undergone Cre mediated recombination. From the data discussed in Chapter 3, the lox 4 line was estimated to contain 18 copies of the BLG transgene within its array. As discussed above from the size of the 10kb RV fragment it is estimated that it contains a single transgene. This means the 'reduced array' represents one copy per cell, while the 'parental array' represents 18 copies per cell. Therefore if the two EcoRV arrays were of equal intensity this would indicate that there were 18 reduced cells to every one unreduced cell in the mammary gland. However, the reduced array was less intense than the unreduced array in figure 4.3. The ratio of unreduced/reduced intensity was calculated using Quantity One Software to be 1.65. This meant that for every one unreduced cell there were 11 reduced

cells. This estimates that 8% of the cells in the mammary gland did not undergo Cre mediated recombination. This is represented in figure 4.5. It was previously estimated that 70-80% of the cells within the lactating mammary gland underwent Cre mediated recombination in the BLG-Cre 74 line (Selbert *et al* 1998). The 20-30% that failed were taken to represent the non-epithelial cells that do not express the BLG-Cre construct.

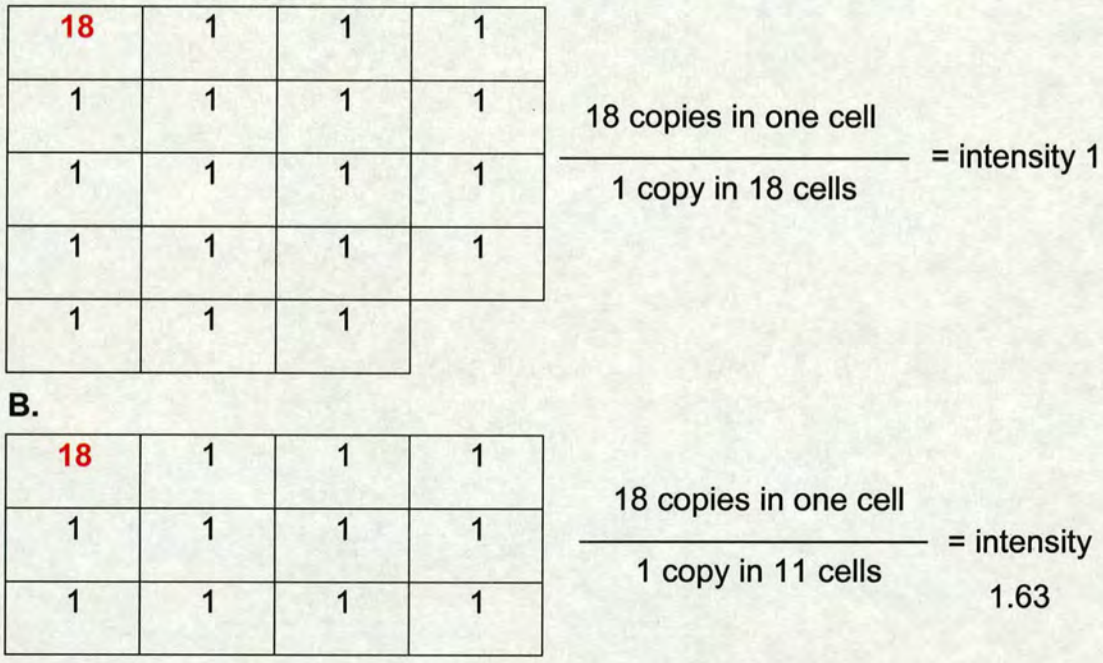


Figure 4.5. Estimate of percentage of cells that underwent Cre mediated recombination. Squares represent one cell, the number inside represents the number of BLG transgenes present. (A) If the ratio of intensity of the unreduced to reduced array was one, then the ratio of cells must be 1 unreduced to 18 reduced. (B) From figure 4.3 the ratio of intensity of the unreduced to reduced array was 1.65, therefore the actual ratio of cells must be 1 unreduced to 11 reduced. The unreduced cells represent 8% of the mammary gland.

However, if the intensities of the fragments in figure 4.3 are considered in light of the various estimates of the cell types of the mammary gland, different figures of lox 4's original copy number are derived. Studies investigating the percentage of each cell type in the mammary gland using immunohistochemistry have estimated that secretory epithelial cells only represent 70% of the total cell population (Sapino *et al* 1990). Therefore only 70% of the gland could express the Cre recombinase transgene. If this figure is used as the number of cells that have undergone recombination to a single copy array, then a much lower estimate of lox 4 copy number is derived:

$$\frac{\text{parental array in 30\% of cells}}{\text{1 copy in 70\% of cells}} = 1.63 \text{ intensity}$$

$$\text{parental array} = \frac{\text{1.63 intensity x 70\% of cells}}{\text{30\% of cells}}$$

$$\text{parental array} = 3.8 \text{ copies}$$

If the same calculation is carried out using an estimate that 80% of mammary cells undergo recombination, the original lox 4 copy number is shown to be 6.52 copies. These figures conflict with the 18 copies estimated by Southern blotting in Chapter 3. It must be considered that these figures rely on the ability to quantify the intensities of each band accurately using a Phosphorimager and its software. As discussed in Chapter 3 a number of issues may result in errors in estimating band intensities by this method. At best, this is a crude method to determine the percentage of cells that have undergone Cre mediated recombination. In the previous studies (Selbert et

al 1990) *in situ* PCR was used to identify cells that had undergone Cre mediated recombination of a reporter construct. This would not have been possible in this case, as no transgenic sequence differences existed between reduced and unreduced cells.

Lox 5 and 9 also showed a reduction in the intensity of the parental array, however no 'reduced array' could be found in double transgenic animals from lox 5 or 9 (figure 4.6).

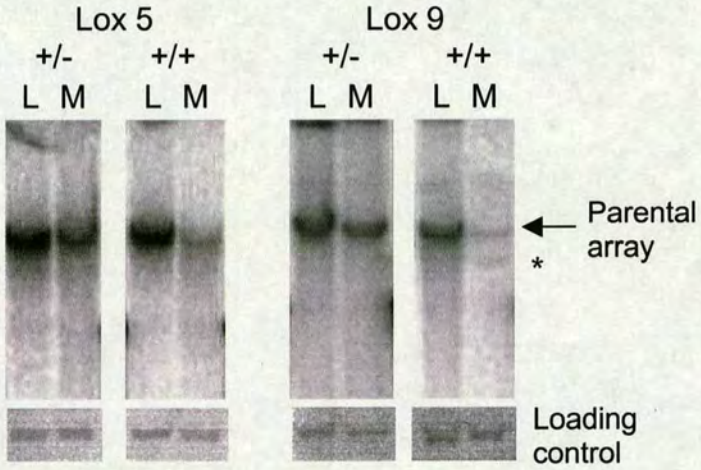


Figure 4.6. EcoRV digestion of lox 5 or 9 double transgenic DNA compared with lox DNA. 10μg liver (L) or mammary (M) DNA was digested with EcoRV, electrophoresed on a 0.8% gel and blotted to Zetaprobe membrane, and then probed with a 1.1kb BLG fragment. +/+ indicates an animal containing both lox 5/Cre or lox 9/Cre transgenes, +/- indicates an animal containing only the lox 5 or 9 transgene. The blot was probed with CFTR probe to control for loading differences. The faint shadow (*) found in the +/+ mammary of lox 9 is not consistent with a true fragment.

If the Cre recombinase deleted different amounts of the array in individual cells, it maybe that the mammary is mosaic for a variety of

different copy numbers. These different arrays would all produce different sized fragments giving no one clear reduced array. This would however, be predicted to give a smear effect in Southern blots due to the presence of multiple fragments of different length. No smears of transgenic fragments were seen for double transgenic animals in either line 5 or 9. An alternative explanation is that in the cells expressing Cre recombinase the BLG-loxP array was completely deleted. This could account for the reduction in the 'parental array' and also the lack of a 'reduced array'. Given that the transgene only contained one loxP site, the complete removal of all transgenic sequence seems unlikely, due to the fact that a loxP site would have to remain in the genome after Cre mediated recombination. It is possible that the explanation for the lack of a 'reduced array' is due to these arrays being 'trapped' within the parental array due to the resolution point of the gel, resulting in the lack of a clear 'reduced' band. This could mean that a 3-4 copy array (i.e 21-28kb fragment) would be co-migrating along with the parental array. It may be possible that discreet 'reduced' arrays for lox5/Cre and lox 9/Cre animals do exist but have not been detected by the methods used. Pulse-field electrophoresis could allow the resolution of 'reduced' but not single copy arrays from the parental array.

Ssp I digestion produces a repeat length fragment as well as junction fragments and was used to estimate the copy numbers of the lines in Chapter 3. Comparisons between liver and mammary lox 4/Cre DNA showed a reduction in the amount of repeat fragment present, indicating a reduction in copy number had occurred (figure 4.7).

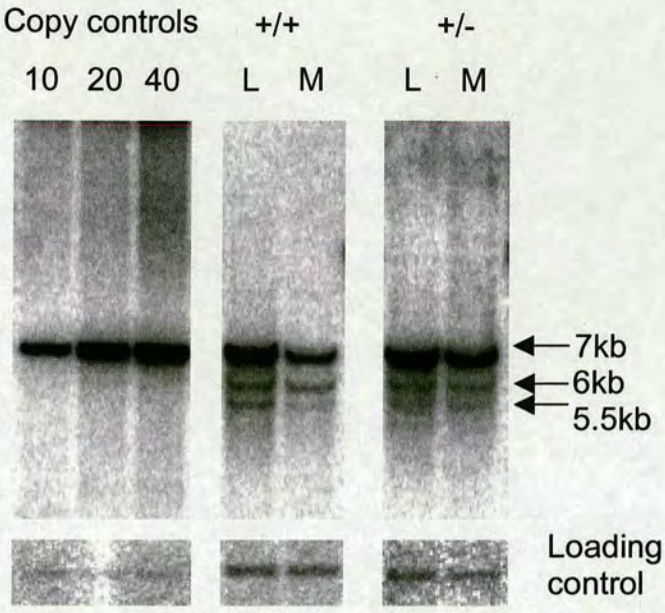


Figure 4.7. SspI digestion of lox 4 double transgenic DNA compared with lox 4 DNA. 10μg liver (L) or mammary (M) DNA was digested with SspI, electrophoresed on a 0.8% gel and blotted to Zetaprobe membrane, and then probed with a 1.1kb BLG fragment. Copy controls were dilutions of the original microinjected 7kb BLG-loxP fragment added to 10μg of non transgenic DNA. +/+ indicates an animal containing both lox 4 and Cre transgenes, +/- indicates an animal containing only the lox 4 transgene. The blot was probed with CFTR probe to control for loading differences. The 6kb and 5.5kb fragments were thought to represent junction fragments. This figure is derived from the same Southern blot as figure 4.3.

SspI digestion was also carried out on the lox 5 and 9 double transgenic animals (figure 4.8). Comparisons between the liver and mammary gland in double transgenic animals shows a reduction of the repeat fragment intensity, suggesting there had been a reduction in copy number in these animals. This is in agreement with the EcoRV digests, which also had a reduction in intensity of the parental array from double transgenic animals.

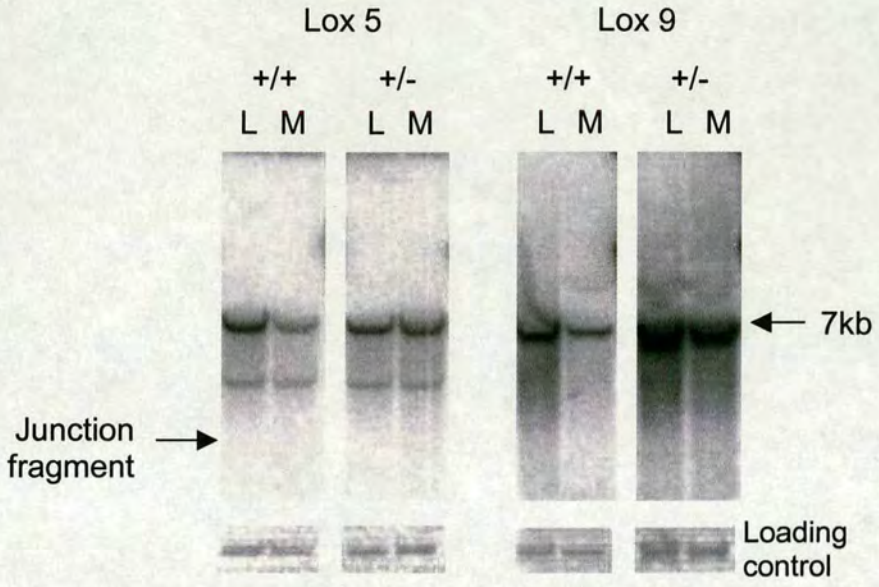


Figure 4.8. SspI digestion of lox 5 or 9 double transgenic DNA compared with lox 5 or lox 9 DNA. 10μg liver (L) or mammary (M) DNA was digested with SspI, electrophoresed on a 0.8% gel and blotted to Zetaprobe membrane, and then probed with a 1.1kb BLG fragment. +/+ indicates an animal containing lox 5/Cre or lox 9/Cre transgenes, +/- indicates an animal containing only the lox 5 or 9 transgene. The blot was probed with CFTR probe to control for loading differences.

As illustrated in figure 4.4, from the size of the 10kb RV locus in the EcoRV digestion it was estimated that there was a reduction to one copy in lox 4/Cre animals. Due to the mosaicism of the mammary DNA it was impossible to carry out more detailed analysis of the transgene fragment, due to the contaminating presence of unreduced copies in the DNA, evident in Southern blots due to its identical sequence. To try and confirm the identity of this fragment the 10kb fragment was probed with a 5′, middle and 3′

probe to determine if it hybridised to all of these (figure 4.9 describes the probes). Figure 4.10 shows that all three probes hybridise to the 10kb fragment indicating it was likely to comprise an intact BLG transgene in addition to its junction fragments The extra bands visible in the double transgenic liver and mammary gland after hybridisation to the 5′ probe are due to this probe hybridising to the promoter of the BLG-Cre construct.

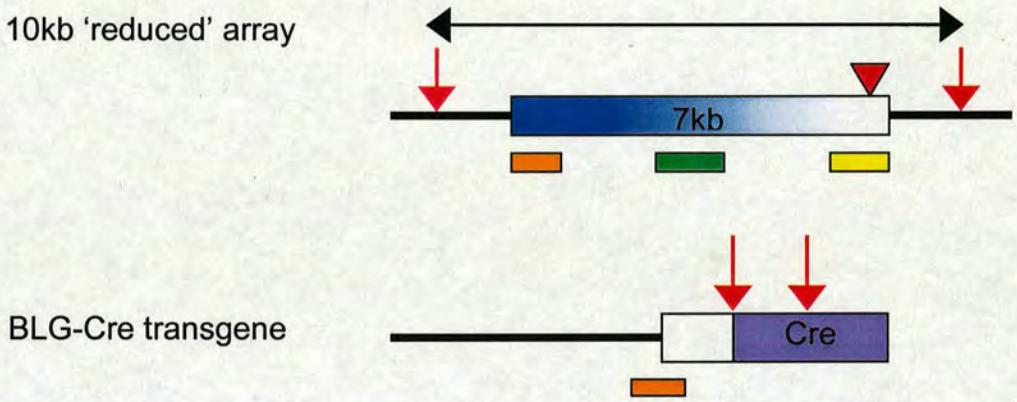


Figure 4.9. Probes and structures of transgenes. If the reduced array contains an intact BLG transgene then all three probes (5' probe orange box, middle probe green box, 3' probe yellow box) would hybridise to the 10kb fragment. EcoRV sites are represented as red arrows, the loxP site is represented by a red triangle. The BLG-Cre transgene contains 4.2kb of the BLG promoter including the first exon (black line and white box). The 5' probe will also hybridise to the BLG-Cre. If the BLG-Cre is a head to tail array the expected band would be approximately 6kb plus any junctions that also contained the 5' region.

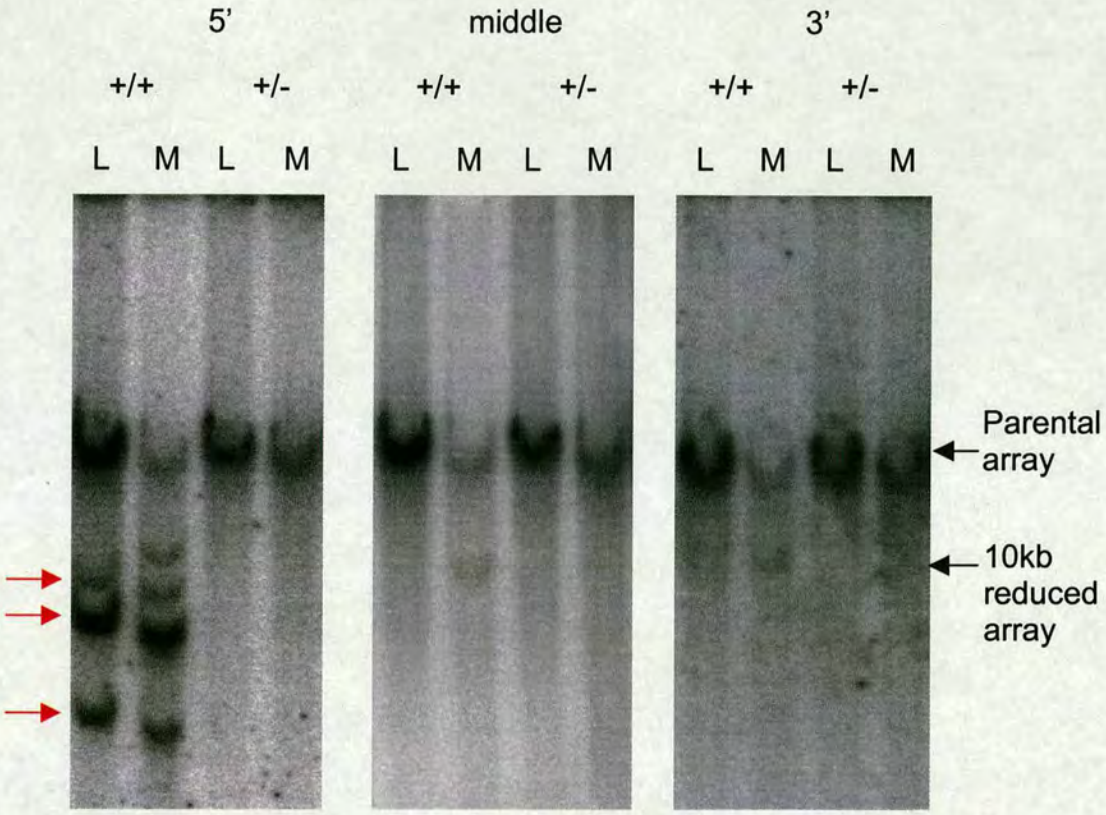


Figure 4.10. Multiple probing of EcoRV digestion of lox 4 double transgenic DNA compared with lox 4. 10μg liver (L) or mammary (M) DNA was digested with EcoRV, electrophoresed and blotted to Zetaprobe membrane, and then sequentially probed with a 5', middle and 3' BLG fragment. +/+ indicates an animal containing both lox 4 and Cre transgenes, +/- indicates an animal containing only the lox 4 transgene. Red arrows indicate fragments (7kb, 6kb, 3kb) generated by the BLG-Cre transgenes.

As far as could be determined by Southern blotting it appears that the lox4/Cre animals do have an intact transgene copy present in the reduced array. Attempts to clarify the results further were all hampered by the contaminating presence of unreduced arrays from cells in the mammary gland that did not express Cre recombinase. Techniques including building a lambda library for screening and cloning out the reduced array as well as

long range PCR were not possible due to the fact that identical copies of the transgene exist in the DNA, making it difficult to distinguish between the two arrays by these means. The Southern blotting cannot however, rule out the possibility of more subtle rearrangements or mutation of the lox 4 'reduced array' which may inactivate the transgene.

By EcoRV digestion there was a reduction in intensity of the parental array of double transgenic lox 5 and 9 animals. However, 5', 3' and middle probes all failed to detect any presence of a 'reduced' fragment in double transgenic lox 5 or 9 animals. There was also no smear effect to indicate multiple transgene arrays of different sizes. However, the use of 0.8% agarose gels may have precluded the identification of 3-4 copy reduced arrays, as these would be trapped within the parental array. The use of this type of gel electrophoresis cannot, therefore, formally exclude the existence of 'reduced' arrays.

4.2.3 Expression Analysis

4.2.3.1 *In situ* hybridisation

Having analysed, as far as possible, the double transgenic animals for copy number reduction, the animals were then analysed for expression. To determine the cellular pattern of the transgene expression after the reduction of copy number in the mammary gland, *in situ* hybridisation was carried out as described in 2.8.

As with the *in situ* analysis discussed in Chapter 3, non-transgenic mammary gland was found to be negative for BLG expression, while expressing β-casein in a uniform manner. Sections from the double lox/Cre transgenics or their lox sibs probed with sense probes had no signal associated with them (as seen in figures 4.13, 4.16, 4.19), indicating non-specific hybridisation did not occur.

Sections from reduced lox 4/Cre animals showed no BLG expression by light or dark field microscopy even after prolonged exposures of up to 6-8 weeks (figure 4.11 and figure 4.12). β-casein was still expressed uniformly as expected, therefore indicating the mammary gland still had the ability to express a milk protein gene in the normal manner. Although no completely positive alveoli was found in any lox 4/Cre animal, very rarely, some apparently partially expressing alveoli were found in lox 4/Cre animals. However, there was no consistent pattern of these dots of expression between sequential slides, it is most likely that these dots represent very low level background, rather than genuine expression. Sibs that contained only the lox transgene showed the parental variegated expression pattern, i.e. approx 50% of the mammary gland expressed. However, as the single lox sibs are of the same generation as the double transgenic animals, this indicated that the silencing in this case was not a progressive silencing of the array over generations, but was due to the presence of the Cre recombinase and therefore the reduced copy number.

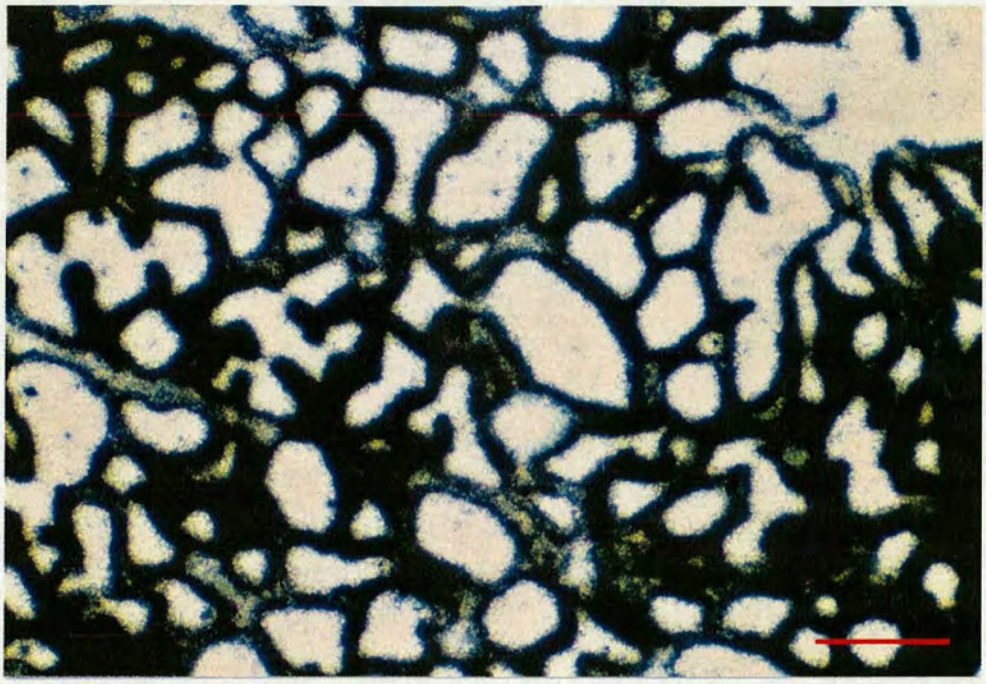
Sections from reduced lox 5/Cre animals showed only small patches of BLG expression (figure 4.14). The single lox 5 sibs mirrored the expression pattern seen in the parental lines, indicating that the reduction in the number

of cells expressing was due to the effects of Cre recombinase. The previously uniform lox 5 line now showed a high number of cells which were silenced. In figure 4.14, some of the alveoli are only weakly positive for β-casein. When photographed by dark field (figure 4.15), all of the alveoli in that section showed expression of β-casein as indicated by the even spread of the black mRNA stain in panel A and the yellow paint of panel C (composite image of dark and light field). This section however, when probed with BLG showed no patches of expression (panel D or F). This apparent variegation of β-casein found by light microscopy may be an artefact of the *in situ* process, due to different areas being more firmly adhered to the slide, washing being different across a slide, or the amount of cells within a given area. Light microscopy is unable to detect lower levels of silver grain s compared with dark field microscopy, and this may also account for the differences seen between figures 4.14 and 4.15.

Sections from reduced lox9/Cre animals showed that only small patches of the mammary gland expressed BLG mRNA (figure 4.17 and figure 4.18). The patches of expression were often quite isolated and only 2-3 patches were seen on any one section. As with the other two lines, when single transgenic lox 9 sibs were analysed they mirrored the parental lines pattern of expression, in this case as variegated expression. As with lox 5 / Cre animals the section of tissue probed with β-casein showed variation between alveoli in the intensity of staining under light microscopy. However, other sections when examined under dark field show overall expression of β-casein as depicted by the yellow paint showing where the reflective silver grains were found in figure 4.18 (panel C).

The *in situ* results did not fit with the hypothesis that the reduction in copy number would produce an increase in expression due to the removal of repeat sequences. These results are also in contrast with published work on the relationship between copy number and variegation (Garrick *et al* 1998). In fact, in all three lines with the expression of Cre recombinase there was a reduction in expression. Lox4/Cre animals had been shown, as far as possible that there was an intact copy of the transgene present. However, the data from the Southern blotting could not rule out that there was small rearrangements or mutations in the transgene which could have functionally inactivated the transgene. When the *in situ* results from double transgenic animals of lox 5 and 9 is considered in light of the Southern blotting data from these animals, the patchy expression may be explained as the deletion of transgenes by Cre recombinase in a variegated manner. The few remaining alveoli expressing BLG may represent cells that have failed to express Cre recombinase, leaving an intact parental array. The non-expressing cells may represent cells where deletion of some copies has produced a 'reduced' array of possibly 3-4 copies which no longer express, and which are not detectable using conventional agarose gel electrophoresis. Indeed the published data on the BLG-Cre line used in these experiments could not formally exclude the possibility that a small percentage of epithelial cells failed to undergo recombination (Selbert *et al* 1998).

A.



B.

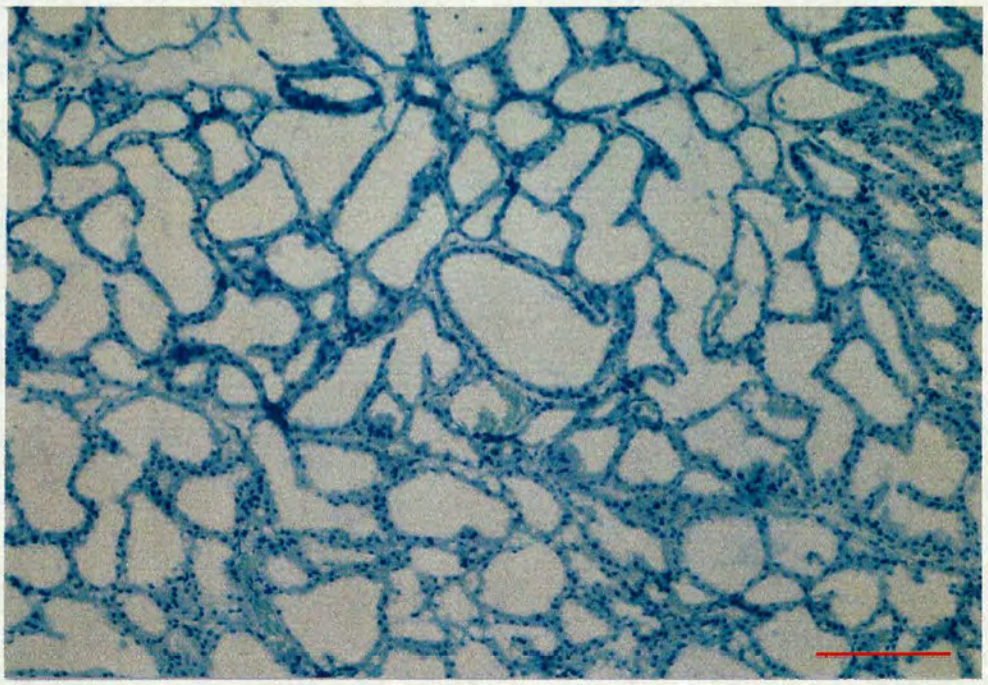


Figure 4.11. *In situ* hybridisation analysis of mRNA expression patterns in lox4/Cre. (A) β-casein, (B) BLG. Photographs were taken after 6 week exposures using a x10 objective. (A) β-casein, (B) BLG. Bar represents 300μm.

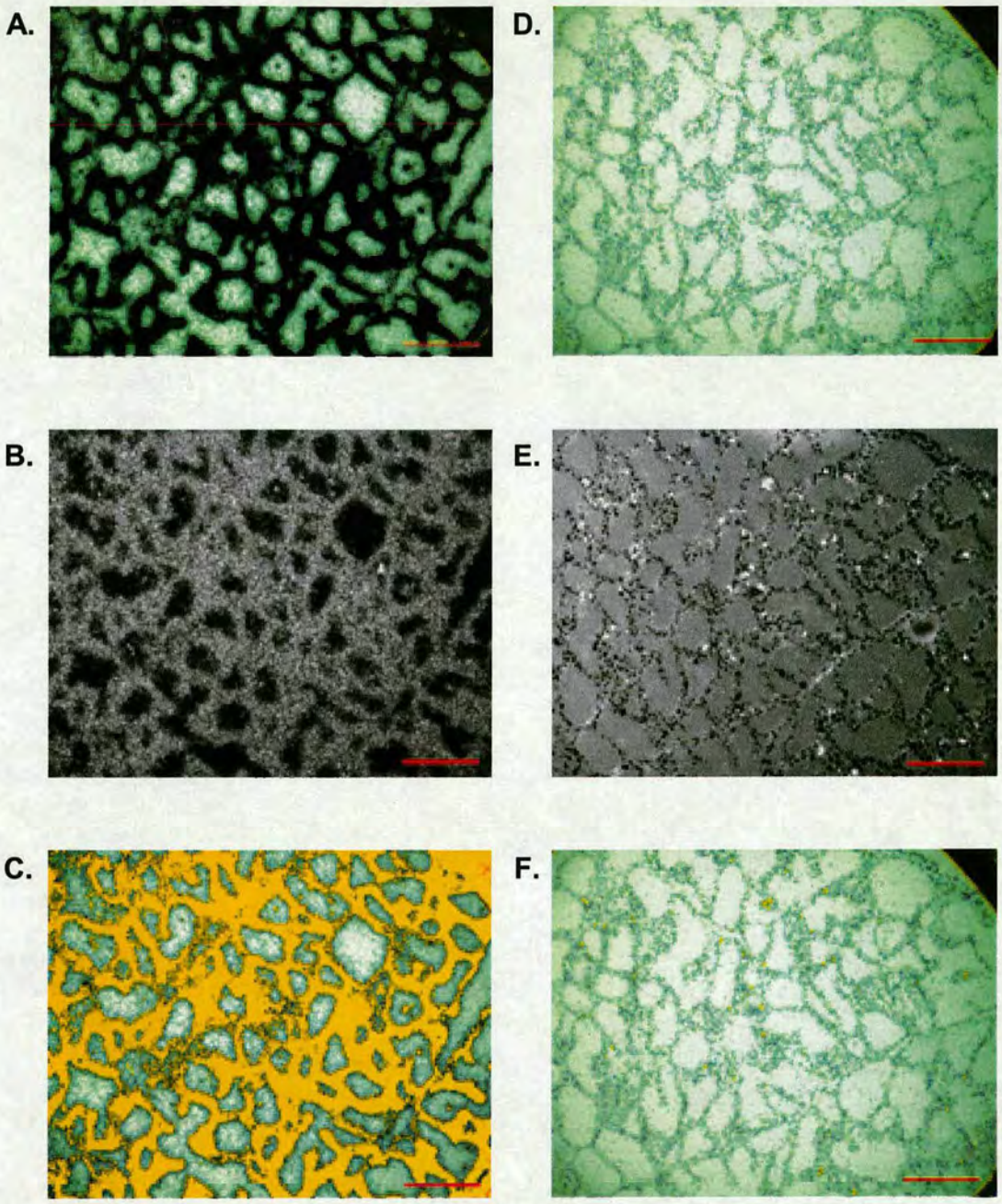


Figure 4.12. Dark field imaging of *in situ* hybridisation of Lox4/Cre. (A,B,C) probed with β-casein. (D,E,F) probed with BLG. Photographs were taken after 6 weeks exposures using a x10 objective. (A) and (D) are light field images. (B) and (D) are dark field images. (C) and (F) are composite images of both light and dark field with the silver grains being represented with a yellow paint. Bar represents 300μm.

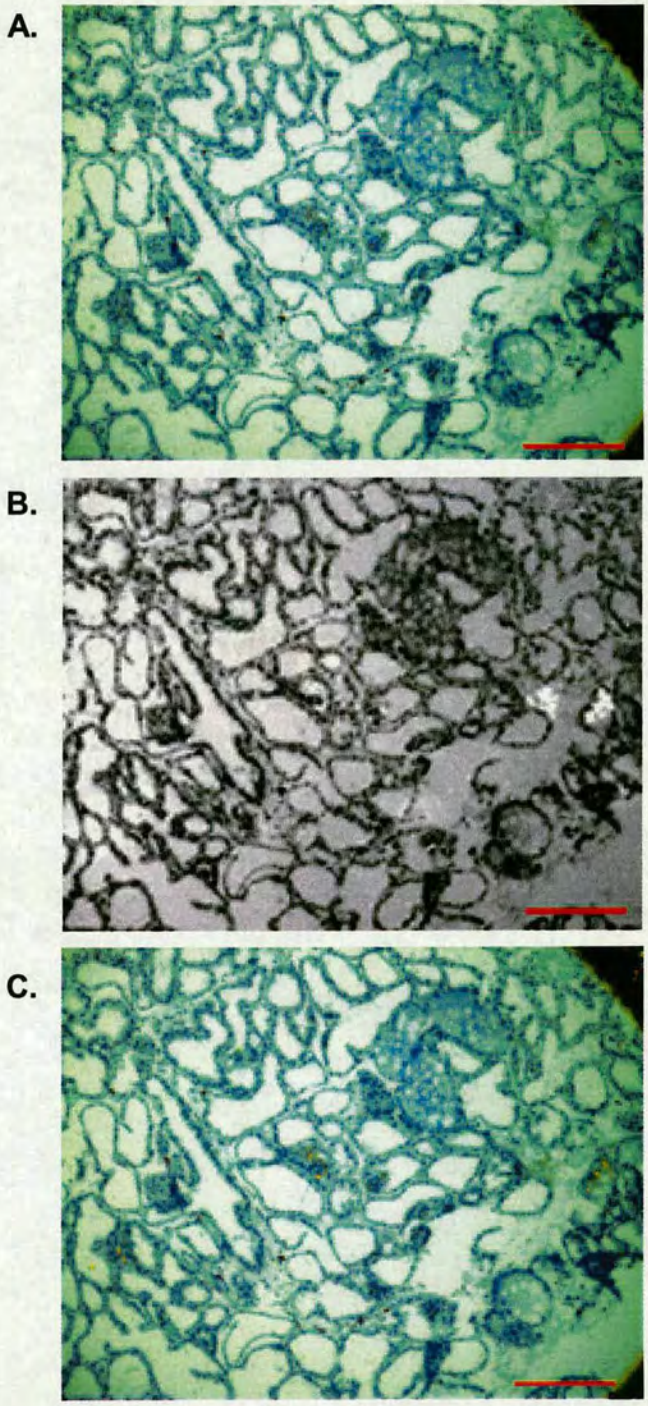
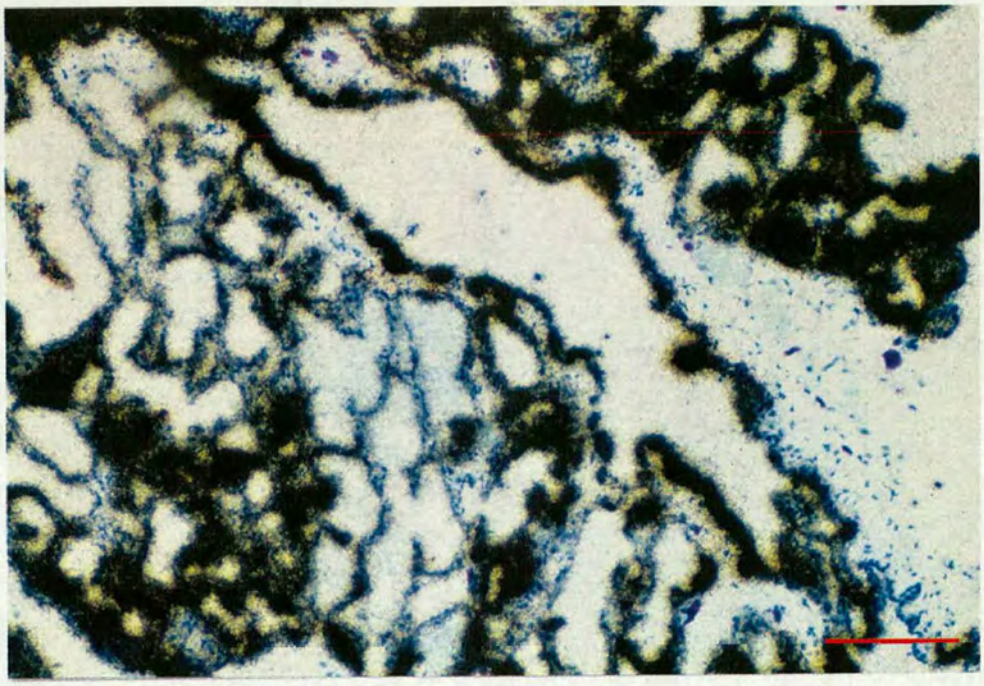


Figure 4.13. Control *in situ* hybridisation of Lox 4/Cre. (A,B,C) probed with sense probe. Photographs were taken after 6 weeks exposures using a x10 objective. (A) is a light field image. (B) is a dark field image. (C) is a composite image of both light and dark field with the silver grains being represented with a yellow paint. Bar represents 300μm.

A.



B.

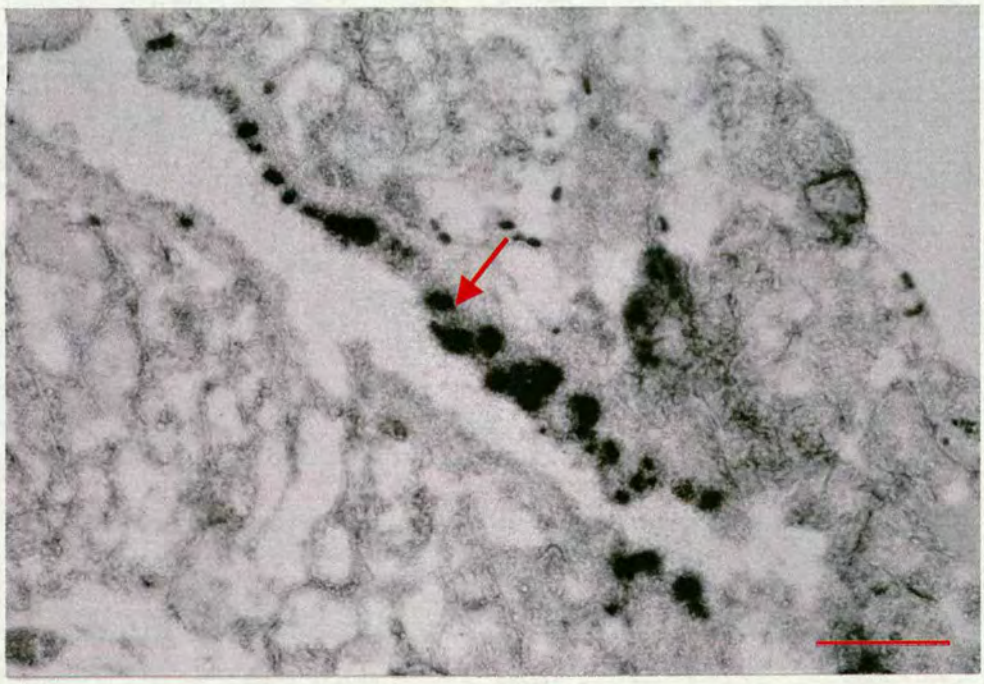


Figure 4.14. *In situ* hybridisation analysis of mRNA expression patterns in lox5/Cre. (A) β-casein, (B) BLG. Photographs were taken after 4 week exposures using a x10 objective. Bar represents 300μm. Red arrow indicates an alveoli with partial expression.

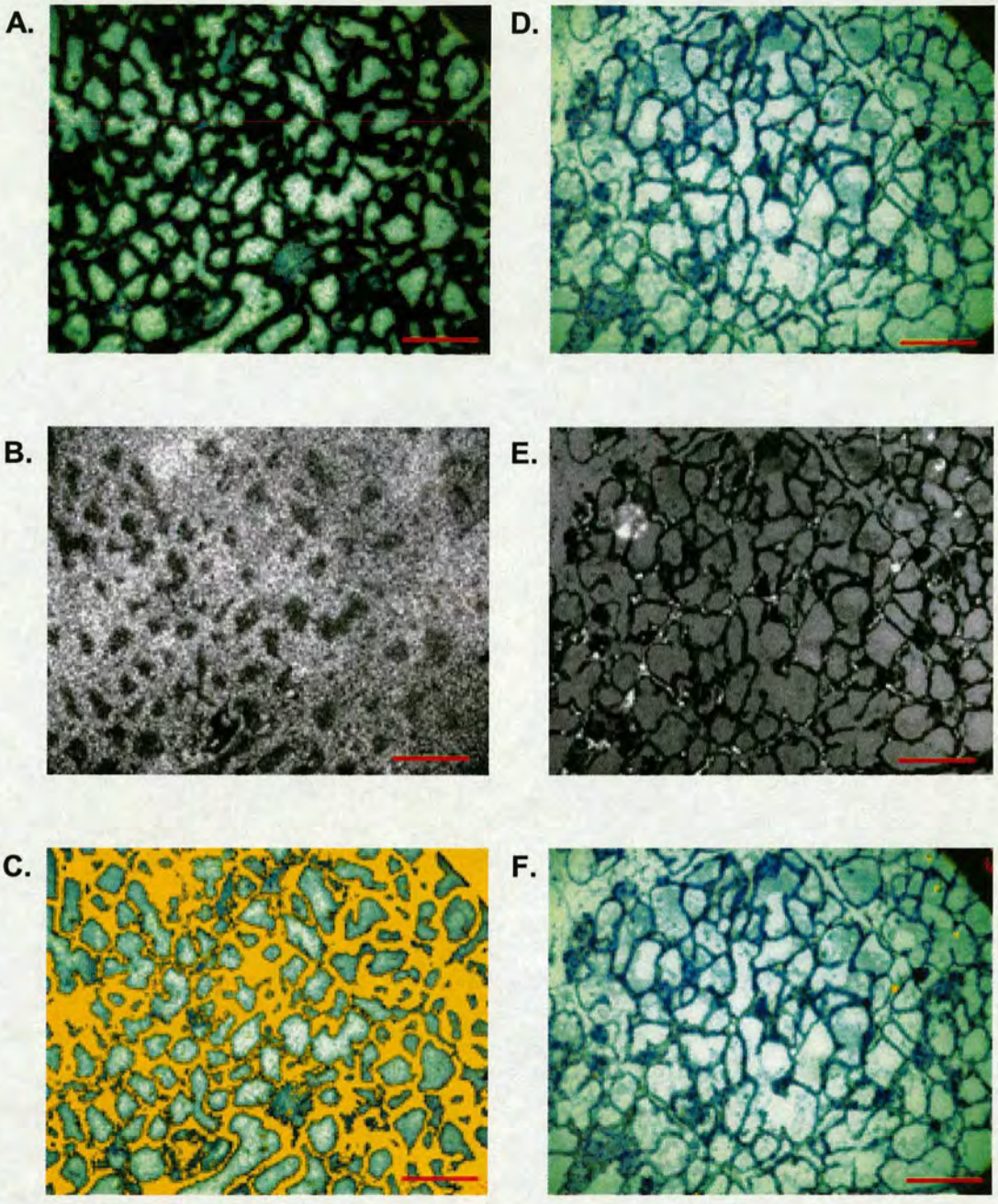


Figure 4.15. Dark field imaging of *in situ* hybridisation of Lox 5/Cre. (A,B,C) probed with β-casein. (D,E,F) probed with BLG. Photographs were taken after 4 weeks exposures using a x10 objective. (A) and (D) are light field images. (B) and (D) are dark field images. (C) and (F) are composite images of both light and dark field with the silver grains being represented with a yellow paint. Bar represents 300μm.

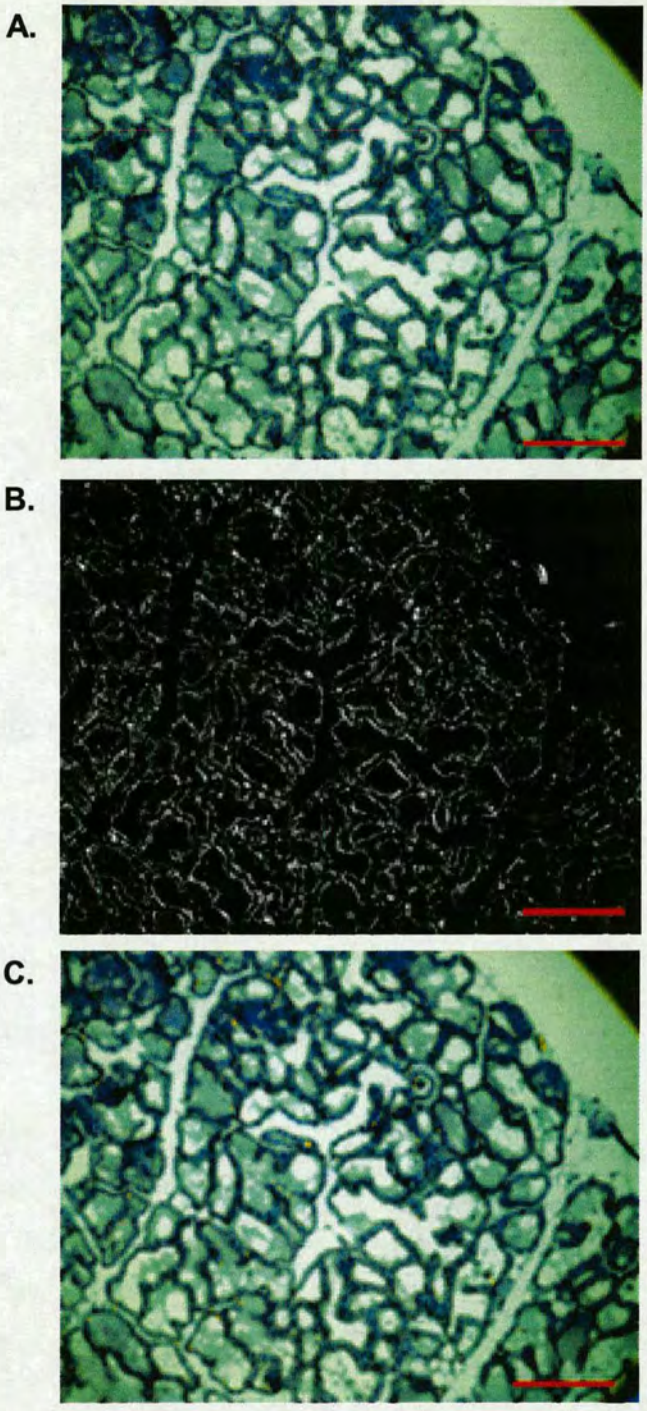
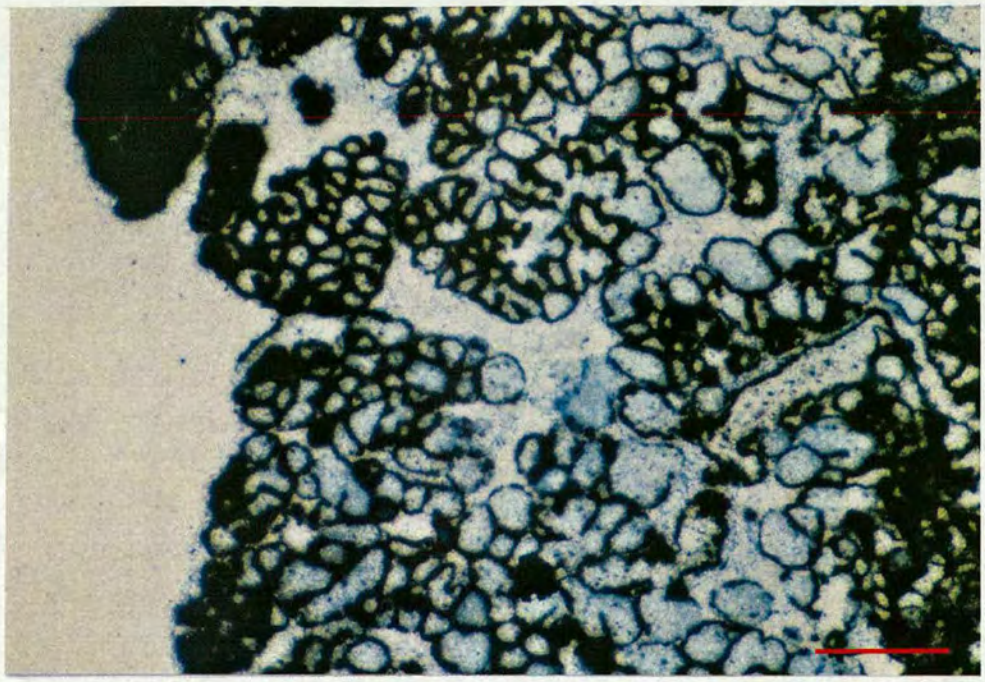


Figure 4.16. Control *in situ* hybridisation of Lox 5/Cre. (A,B,C) probed with sense probe. Photographs were taken after 4 weeks exposures using a x10 objective. (A) is a light field image. (B) is a dark field image. (C) is a composite image of both light and dark field with the silver grains being represented with a yellow paint. Bar represents 300μm.

A.



B.

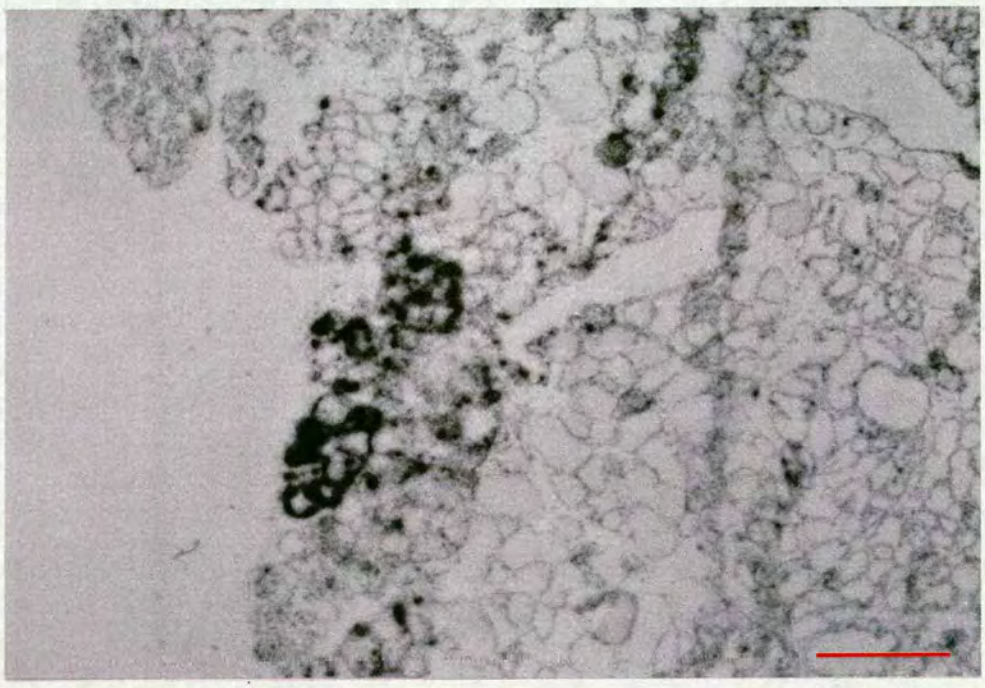


Figure 4.17. *In situ* hybridisation analysis of mRNA expression patterns in lox9/Cre. (A) β-casein, (B) BLG. Photographs were taken after 4 week exposures using a x4 objective. Bar represents 750μm.

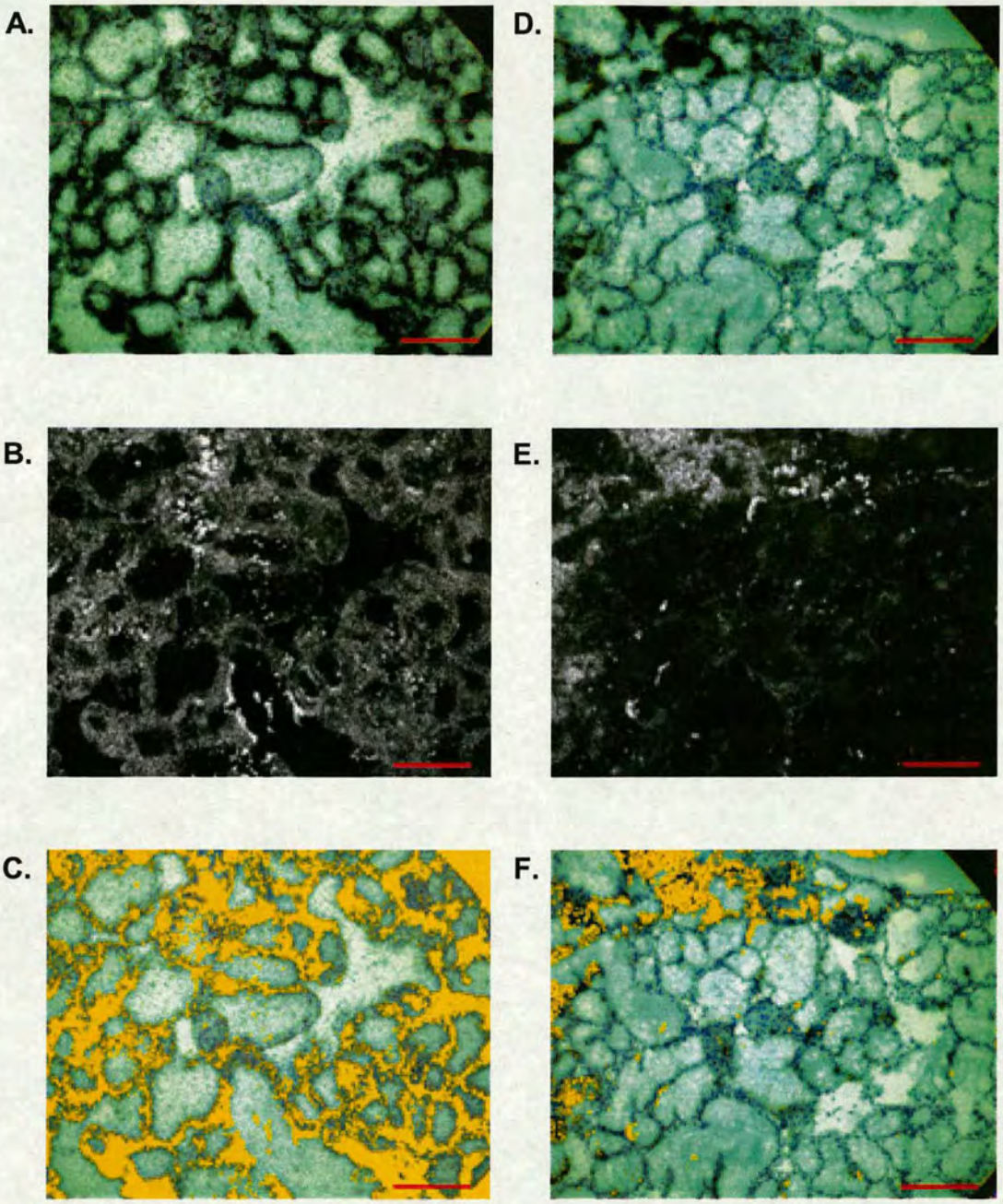


Figure 4.18. Dark field imaging of *in situ* hybridisation of Lox 9/Cre. (A,B,C) probed with β-casein. (D,E,F) probed with BLG. Photographs were taken after 4 weeks exposures using a x10 objective. (A) and (D) are light field images. (B) and (D) are dark field images. (C) and (F) are composite images of both light and dark field with the silver grains being represented with a yellow paint. Bar represents 300μm.

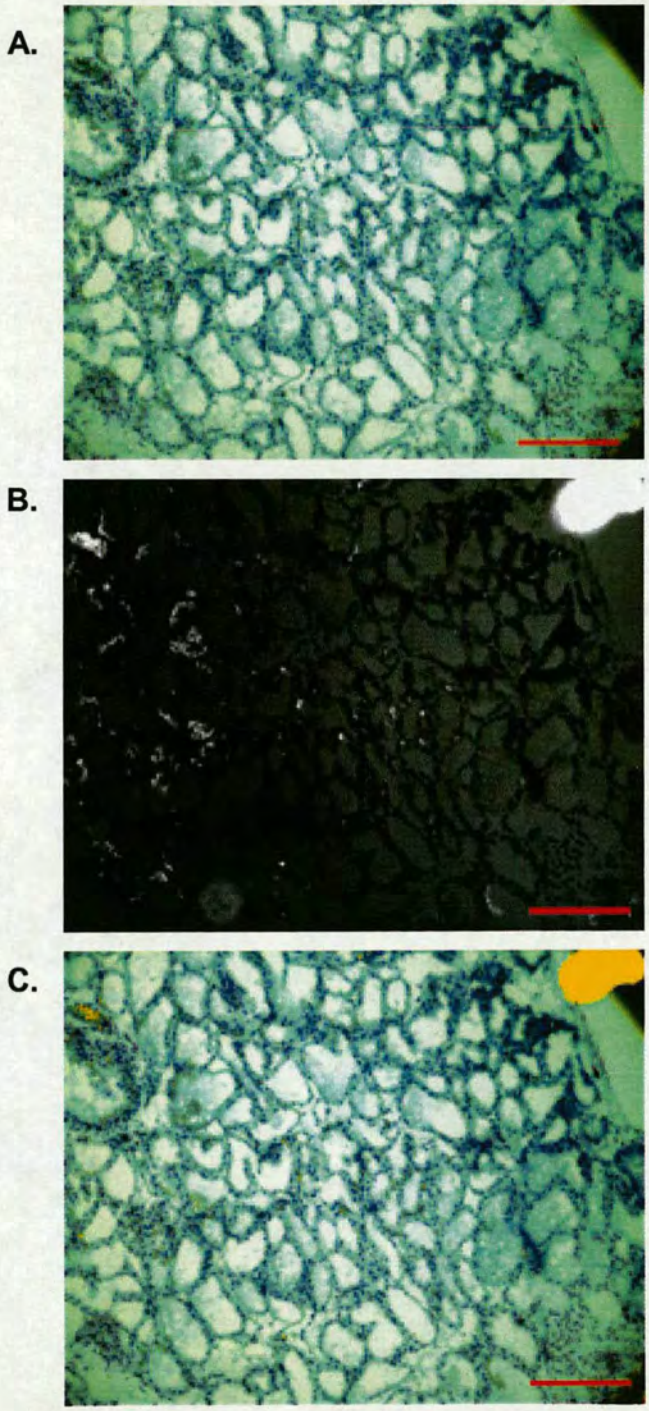


Figure 4.19. Control *in situ* hybridisation of Lox 9/Cre. (A,B,C) probed with sense probe. Photographs were taken after 4 weeks exposures using a x10 objective. (A) is a light field image. (B) is a dark field image. (C) is a composite image of both light and dark field with the silver grains being represented with a yellow paint. Bar represents 300μm.

As with the *in situ* analysis in Chapter 3, the amount of cells expressing BLG was quantified. Alveoli were scored according to the amount of area seen to be expressing BLG RNA; negative when no cells in the alveolus expressed, -/+ when there was partial expression from an alveolus, and positive when all the cells in an alveolus expressed. Figures giving the percentages of alveoli in each group are given in table 4.1. A dotplot of the combined positive alveoli for the double transgenic animals is compared to values for the parental lox lines (figure 4.20).

Animal ID	*Line*	*Negative alveoli*	*Partly positive alveoli*	*Positive alveoli*
GB 7.1	Lox 4/Cre	100	0	0
GB 7.4	Lox 4/Cre	100	0	0
GB 6.7	Lox 4/Cre	99.9	0.1	0
GB 6.11	Lox 4/Cre	98.9	1.1	0
GB 7.8	Lox 4/Cre	98.9	1.1	0
GB 2.14	Lox 5/Cre	98.9	0.9	0.2
GB 2.8	Lox 5/Cre	95.6	3.8	0.6
GB 2.15	Lox 5/Cre	95.0	5.0	0
GB 2.1	Lox 5/Cre	94.3	5.6	0.1
GB 2.3	Lox 5/Cre	81.8	18.1	0.1
GB 11.7	Lox 9/Cre	97.7	2.3	0
GB 11.6	Lox 9/Cre	97.6	2.4	0
GB 9.1	Lox 9/Cre	88.6	10.8	0.6
GB 11.1	Lox 9/Cre	82.6	14.9	2.5
GB 11.2	Lox 9/Cre	79.3	19.4	1.3

Table 4.1. BLG expression values of double transgenic animals by *in situ* hybridisation. Ten fields per section and two sections for each animal were counted to derive a percentage for the three categories.

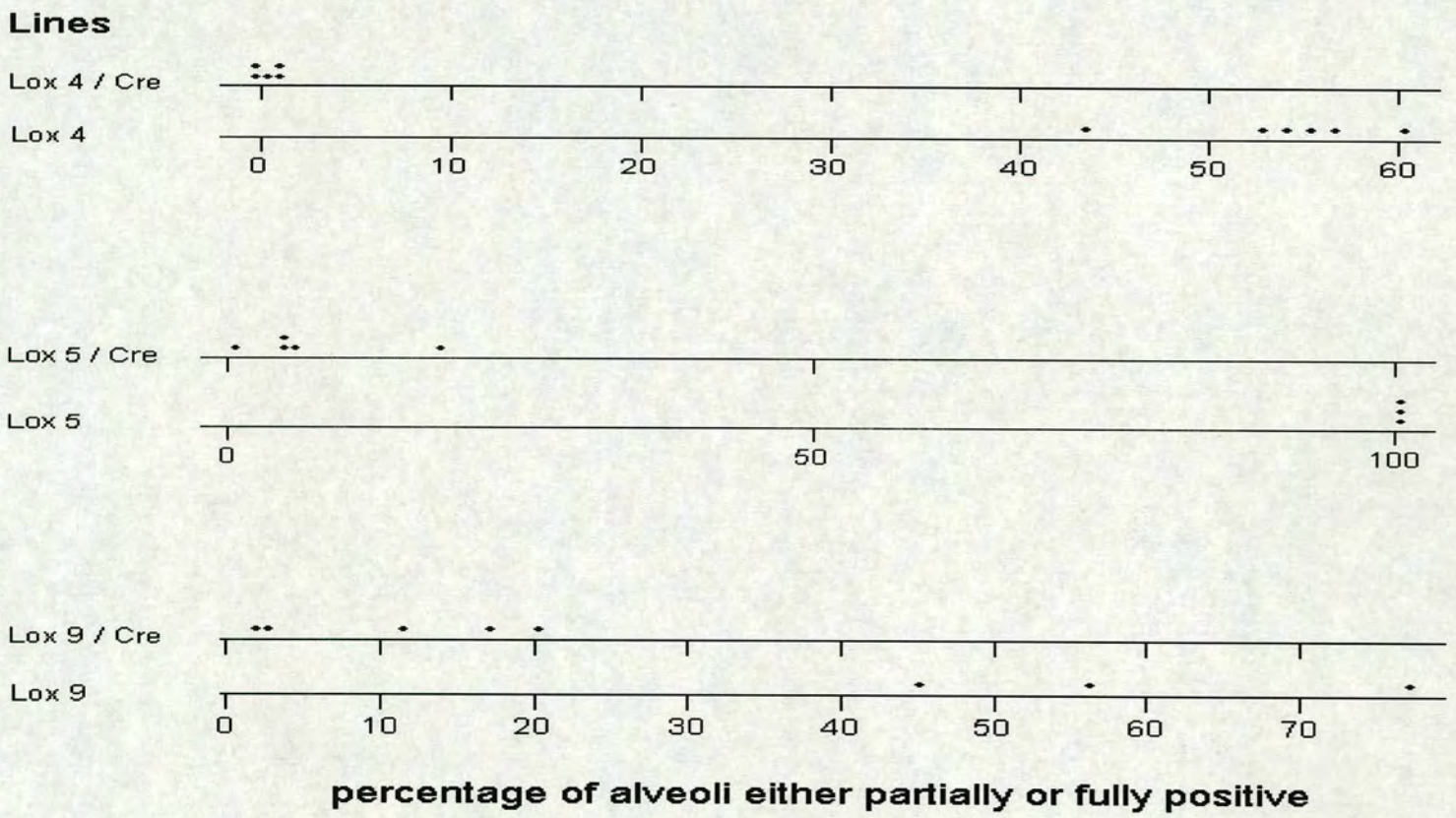


Figure 4.20. Dotplot comparing double transgenic animals with the parental lox lines. Each dot represents an animal. The percentage alveoli that were counted partial or fully positive were combined to give an overall view of the total expression from the mammary gland.

The counts were analysed using GenStat software to produce Table 4.2 combining the values of alveoli counted as partly positive and those fully positive. Standard deviations (s.d.) and coefficient of variance are both measures of the amount of variablitity within a group i.e. the differences between individuals from the same transgenic line.

Line	*Range of percentages of positive alveoli*	*n*	*Mean*	*s.d*	*c.v*
Lox 4 / Cre	0-1.1	5	0.5	0.6	124.3
Lox 5 / Cre	1.1-18.2	5	6.9	6.6	95.8
Lox 9 / Cre	2.3-20.7	5	10.9	8.4	77.6

Table 4.2. Statistical analysis of double transgenic *in situ* counts. Range is given as a percentage of the total alveoli which had some expression, *n* = number of animals per line, *s.d.* standard deviation, *c.v.* = 100 x s.d. / mean.

It can be seen from table 4.2 that all the lines now show variation in expression between cells. Lox 5 had previously been uniform but after Cre recombinase expression, its mean value had dropped from 100% to 6.9% and had a high level of variation between individuals within the line. The percentage of expression in lox 5/Cre is similar to the percentage of cells that were calculated to have failed to undergo recombination in lox 4/Cre (figure 4.5). This supports the theory that the patches of expression seen in lox 5/Cre and lox9/Cre are due to unrecombined cells, which have retained the parental array and therefore express BLG. The other cells, although no 'reduced' array was found, may not express from a possible 3-4 copy array,

leading to the patches of expression from unrecombined cells next to 'reduced' but silent cells.

Statistical analysis was carried out between the parental lox lines and the double transgenic lines to determine if there was a statistical difference between them (table 4.3). It is clear that there is a statistical difference in percentage of expressing alveoli between the parental and reduced lines in all three cases, with the lox4/Cre group showing the highest significance due to its large decrease in expression.

Comparison	*Statistical significance P value*
Lox 4 to Lox 4/Cre	0.006
Lox 5 to Lox5/Cre	0.036
Lox 9 to Lox 9/Cre	0.036

Table 4.3. Statistical probability analysis of parental and reduced lines. The two lines (parental and reduced) were compared for the percentage of total alveoli counted as positive using the Mann-Whitney test (a non-parametric test) to calculate the statistical probability.

4.2.3.2 Northern blots analysis

Northern blots were carried out to determine if the levels of expression shown by *in situ* analysis was mirrored in total RNA isolated from

mid-lactation mammary tissue. However, no BLG RNA was detected from the double transgenic animals even after long exposures. It must be considered that the *in situ* hybridisations had to be exposed for upto 8 weeks to produce a signal and it is therefore unlikely that northern blots would be sensitive enough for the amount of expression found in the reduced animals.

4.2.3.3 Milk protein analysis

SDS-PAGE analysis was used to determine if BLG protein was present in the double transgenic animals. Mouse milk samples were prepared as described (see 2.7.1) and were loaded at 1/200 or 1/100 dilutions along side ovine BLG standards. Single transgenic sibs showed levels of protein comparable to the animals analysed in chapter 3. Double transgenic animals from lox 4, 5 and 9 showed no detectable levels of protein by Coomassie staining (figure 4.21 shows lox 9/Cre animals).

Western blots were carried out to determine if any detectable protein was present in the milk samples from the reduced animals. The detection level of Western blotting is approximately 1-10ng of protein, compared to 0.1 – 0.5μg per track with Coomassie Blue staining (Harlow and Lane 1988). Western blotting, therefore, has a higher sensitivity level than staining SDS-PAGE gels.

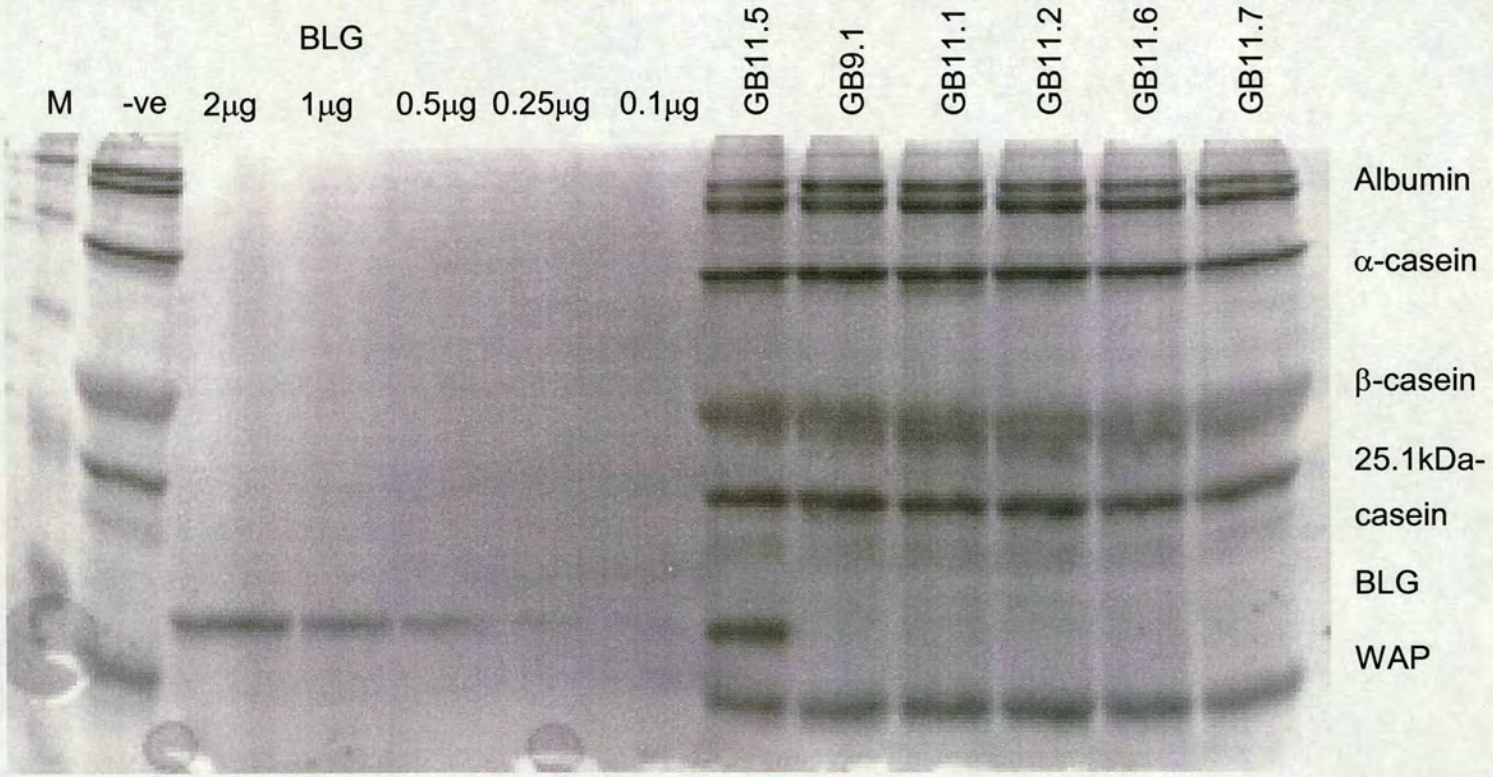


Figure 4.21. SDS-PAGE analysis of lox9/Cre milk proteins. Mouse milk samples were electrophoresed in 18% polyacrylamide gels under reducing conditions and stained with Coomassie Blue. M indicates Rainbow marker (BioRad). -ve indicates a 1/100 dilution of defatted control non-transgenic mouse milk sample. 1/100 dilutions of defatted milk from transgenic animals from lox 9/Cre animals and a lox 9 only animal (GB11.5) was loaded along side standards of BLG.

When milk samples from reduced animals were Western blotted there was a band present in only one reduced lox9/cre animal (GB9.1), but the signal was substantially weaker than the 0.1μg standard (figure 4.23). However, it must be considered that this was a 1/50 dilution, as compared to Westerns of the parental lox lines (figure 3.33) which were 1/200 dilutions. It is of note that GB11.1 had more positive cells by quantification of *in situ*, but it failed to produce a band by Western blotting. The absence of bands from the other animals shown to have limited expression by *in situ*, draws into question the validity of the band of GB 9.1. It may be that this is a false positive, however to confirm this more blotting would be required. The Western blotting and SDS-PAGE results both indicate that the reduction in the percentage of cells expressing has led to a dramatic decrease or complete abolishment of production of BLG protein

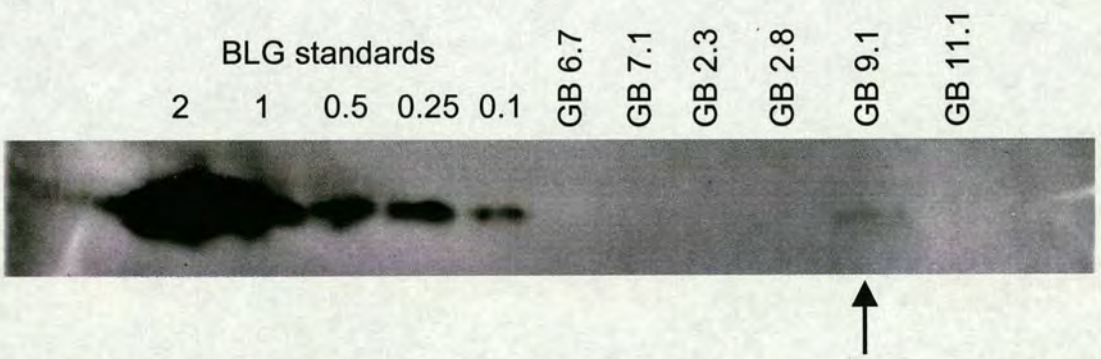


Figure 4.22. Western blot of milk proteins from double transgenic animals. 1/50 dilutions of milk protein samples were run on a SDS-PAGE gel, electroblotted to nitrocellulose and probed with rabbit anti – BLG and then HRP conjugated anti rabbit IgG. The BLG standards are in μg. The putative positive lane is indicated by a black arrow.

4.3 CONCLUSIONS

The use of a BLG driven Cre recombinase resulted in double transgenic animals that had reductions in copy number of the BLG-loxP transgene. This was however, a mosaic pattern of deletion as some of the cells in the mammary gland do not express the Cre transgene. Previous studies have shown that the BLG-Cre is restricted to mammary epithelial cells (Selbert *et al* 1998). The breeding strategy used to cross BLG-loxP and BLG-Cre animals had resulted in a decrease in intensity of the parental array in all three lines. This was specific to the mammary gland and did not occur in sibs that contained only the BLG-loxP transgene.

All the lines showed a decrease in intensity of the parental array, however only one line lox 4, contained a distinct fragment consistent with the presence of a reduced array containing an intact BLG transgene. Multiple probing of Southern blots indicated that it hybridised to both promoter sequence and parts of the coding sequence, suggesting that it was a single copy intact transgene. However, the Southern blotting data is unable to formally exclude the possibility of more subtle rearrangements or mutations which may have functionally inactivated the transgene.

In the other two lines lox 5 and 9, although they showed a decrease in intensity of the parental array, no fragment consistent with a single copy transgene was found in any of the these double transgenic animals. One possibility was that the Cre recombinase was producing different sized reduced fragments in different cells. However, the Southern blotting data

does not show a smear to indicate that this is the case. One interpretation is that a complete deletion of the BLG-loxP array has occurred in Cre expressing cells, which would account for the decrease in intensity of the 'parental array' and the lack of a 'reduced array'. Alternatively the lox 5 and 9 animals may have undergone limited deletion to 3-4 copies. These would not have been identifiable from the Southern blots used in this chapter due to the resolution point of the agarose gels.

The hypothesis for this project was that high copy number arrays often variegate and that a reduction in their copy number should lead to uniform and higher expression. This was not the case. The reduction in copy numbers led to a reduction in the number of cells expressing BLG mRNA at the *in situ* level. In fact with less copies present, a previously uniform line now expressed in a variegated manner. This was unexpected and was contrary to published work on reducing copy numbers (Garrick *et al* 1998). The reduction in copy number also correlated to a decrease in the level of protein whereby no protein was detectable by Coomaisse staining of SDS-PAGE. Only one reduced animal showed expression by Western blotting. However, given that animals with comparable expression by *in situ* hybridisation did not show expression by Western blotting, it seems possible that this is an artefact and not a true positive.

There are a number of possible explainations for this increase in silenced cells which can be thought of as technical or biological reasons for the result.

Technical explanations include the possibility that the transgenes have been rearranged. In lox 4/Cre, no expression was found. Lox 4/Cre was the only line that had a demonstrated fragment consistent with the presence of a single intact BLG transgene. The Southern blotting data could not however, formally prove the existence of a functional transgene within the 10kb RV locus. Thus it is a possibility that this 10kb RV locus has been rearranged or mutated.

In lox 5 and 9 lines after Cre expression there was no data to support the existence of a 'reduced array'. It is therefore possible that in these animals there has been complete deletion or partial deletion of the transgene to a small array. Without the use of PFGE to resolve fragments of 20- 40kb, it cannot be stated that small reduced arrays do not exist in these cells . The areas of BLG expression could be explained as a failure to undergo Cre mediated recombination in these cells. This would mean that the expression of the BLG-Cre itself had been variegated. The BLG-Cre 74 line had been analysed previously, and had shown uniform Cre expression. However, the possibility that some of the epithelial cells did not undergo recombination was not formally exclude (Selbert *et al* 1998). No *in situ* hybridisation on the Cre expression pattern of the double transgenic animals was carried out in this study. Previous immunohistochemistry provide unreliable in detecting Cre expression (Selbert et al 1998), and as the use of *in situ* PCR would not have been informative in this study it was not carried out.

Another technical reason could be that the reduction had removed only the active copies. There is no published evidence on how Cre recombinase removes sequences between multiple loxP sites; i.e. sequentially

or by removing the DNA between the extreme loxP sites in one reaction, so it may be possible that only active copies are removed. However, the patches of expression in lox5/Cre and lox9/Cre animals would mean that in some cases active copies were left behind, while in other cells the active copies had been rearranged. This would mean that were different recombinations occurring in different cells.

If the reason for the decrease in expression is not a technical explanation, 'biological' reasons may be involved. The first model suggests that reduction in the mammary gland is too late developmentally to rescue the transgene silencing. This explanation assumes that some sort of 'mark' is placed on the transgene and once this has occurred, alteration of the structure (i.e. the copy number) will have no effect. The patterns of expression by *in situ* suggests that the patches could be clonal in origin. This may suggest that the decision to silence the transgene is an early one. It must be considered that the mammary gland is only fully differentiated after pregnancy in adult females. BLG transgenes, including the BLG-Cre transgene, have been shown to be regulated coordinately with endogenous mouse milk genes, with their expression increasing through pregnancy to reach the highest level of expression during lactation (Harris *et al* 1991, Selbert *et al* 1998). Hence, deletion at a late stage, i.e. when the BLG-Cre transgene is activated, may not overcome silencing.

A second biological explanation is that BLG copies provide a 'buffer' effect, insulating a number of copies from genome silencing effects. If the position within the genome was suppressive to expression of the transgenes and this occurred by the spread of heterochromatin, then the large arrays

may allow some internal copies to escape this spread. If there were less copies then the genome effects would spread through the whole array, silencing the single copies left. Therefore the reduction in copy number would actually increase the number of cells that are silenced.

CHAPTER FIVE

REDUCTION OF TRANSGENIC COPY NUMBER BY MICROINJECTION

5.1 INTRODUCTION

In Chapter 4, high copy number BLG-loxP lines underwent Cre catalysed recombination in the mammary gland and the consequences for transgene expression were analysed. After reduction of copy number, the expression levels were lower, and the percentage of the gland that expressed had considerably decreased.

Microinjections of Cre recombinase constructs or mRNA have both been used successfully to produce early recombinations in embryos (Araki *et al* 1995, de Wit *et al* 1998). This method has an advantage over deletion in the mammary gland, in that all the cells within the animal should carry the reduced transgene array, allowing for clearer analysis without the contaminating presence of unreduced arrays from non-Cre expressing cells. The other main advantage of this is that it allows the establishment of reduced lines, allowing larger numbers of animals to be investigated. In addition, the mammary gland is a late developmental organ, and it is only fully developed in the adult female after pregnancy. It differentiates in response to a number of factors present during pregnancy and lactation (Hennighausen and Robinson 1998). Given that the BLG-Cre recombinase

expression profile is co-ordinately expressed with other milk proteins, the Cre expression only occurs at substantial levels in the adult mammary gland during lactation (Selbert *et al* 1998). One possible explanation for the results in Chapter 4 was the reduction in copy number had occurred beyond a developmental checkpoint, which would allow the 'rescue' of expression from the reduced arrays. Work in this chapter was meant to create a reduction in copy number at a much earlier developmental time point, and to investigate the consequences of this modification on the expression pattern after reduction. This would be a reduction before differentiation of the mammary gland, which contrasts with the reduction in Chapter 4, which was after the tissue had started to differentiate.

Phosphoglycerate kinase (PGK) is expressed in all somatic cells and has been used to drive early and uniform Cre recombinase expression (Adra *et al* 1987, Lallemand *et al* 1998). PGK-Cre was therefore chosen to produce a Cre catalysed reduction early in development by microinjection of a circular plasmid carrying the PGK-Cre construct.

Microinjection of a circular plasmid may reduce the chances of it integrating into the genome. Continued Cre recombinase expression may be detrimental, as there are now indications that in some situations Cre recombinase may target pseudoloxP sites within the genome, leading to illegitimate recombination events to occur (Thyagarajan *et al* 2000, Schmidt *et al* 2000). This method therefore avoids any complications from continued Cre expression.

The three lines discussed in Chapter 4 were used for microinjection of a PGK-Cre construct; lox 4 and 9 which had variegated expression and lox 5 which had uniform expression. Microinjected transgenic animals were analysed for changes in copy number by Southern blotting and lines were established. Expression analysis was carried out as before using *in situ* hybridisation, Northerns and milk protein analysis. The results are discussed below.

5.2 RESULTS

5.2.1 Analysis of copy number reduction

Founders from the microinjection were analysed for the production of a 'reduced array' by EcoRV digestion, which cannot cut within the BLG-loxP transgene (similar to the situation in Chapter 4). A number of animals from the lox 4/PGK Cre line (termed PLC A) showed an EcoRV fragment, which when compared with lambda fragments of known size was estimated to be 10kb in size (figure 5.1). This fragment co-migrated with the 10kb RV locus from Chapter 4 lox4/Cre animals. Some animals also contained longer fragments by EcoRV digestion (i.e. PLC A58; an estimated 17kb fragment). These animals were therefore likely to be mosaic for different copy number 'reduced arrays'. This was probably due to Cre mediated recombination occurring after the embryo had started to divide. The larger fragments were thought to represent a 'reduced array' of more than one full transgene plus its junction fragments, due to its size. The 10kb fragment is referred to as the 10kb RV locus, while the 17kb fragment is referred to as the 17kb RV locus.

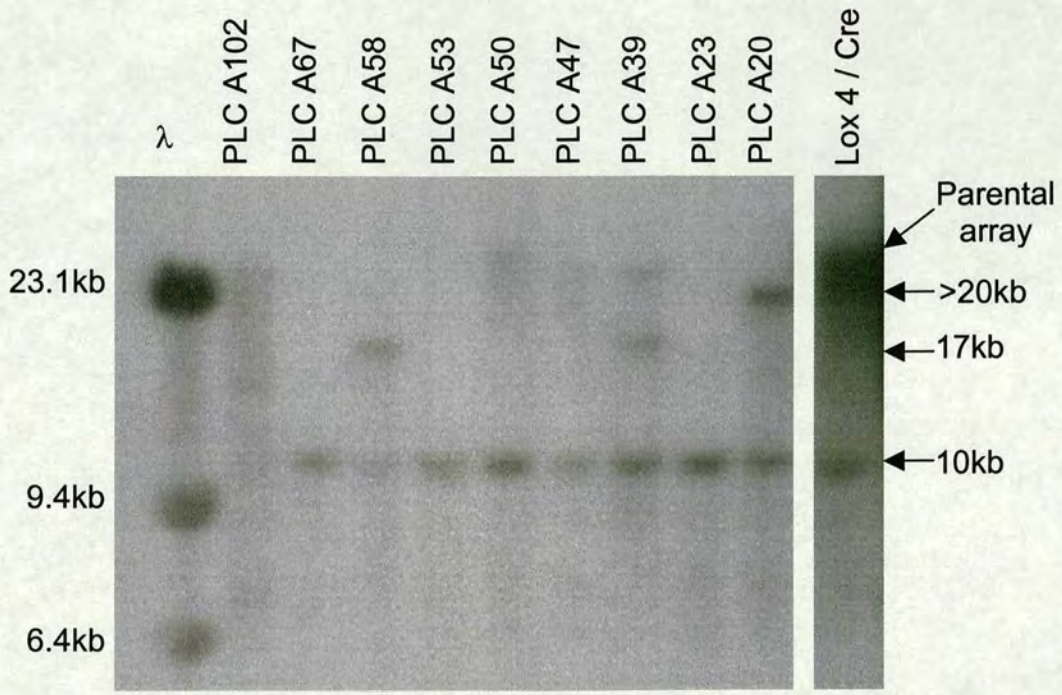


Figure 5.1. EcoRV digestion of microinjected PLC A founder mice. 10μg tail DNA from PLC A founders was digested with EcoRV, electrophoresed on a 0.7% gel, blotted to Zetaprobe and hybridised to a 1.1kb fragment of BLG (described in Chapter 2). Lox4/Cre DNA also produced a 10kb fragment as described in Chapter 4. λ indicates a lambda HindIII ladder with stated sizes.

Four of these positive animals were used to establish lines; PLC A67, PLC A58 which was mosaic for two arrays, PLC A53 and PLC A23. The multiple arrays of PLC A58 segregated upon breeding to the next generation as shown later in figure 5.4. The ratio of segregations is shown in table 5.1.

Animal	*Generation*	*Sex*	*10 kb EcoRV*	*17kb EcoRV*
PLC A58	Founder	M	Yes	Yes
PLC A58.2	G1	M	No	Yes
PLC A58.3	G1	M	No	Yes
PLC A58.4	G1	M	No	Yes
PLC A58.5	G1	M	No	Yes
PLC A58.6	G1	M	No	Yes
PLC A58.8	G1	M	No	Yes
PLC A58.11	G1	F	No	Yes
PLC A58.12	G1	F	No	Yes
PLC A58.13	G1	F	Yes	No
PLC A58.15	G1	F	No	Yes
PLC A58.17	G1	F	No	Yes
PLCA58.18	G1	F	No	Yes
PLC A58.19	G1	F	No	Yes
PLC A58.23	G1	M	No	Yes
PLC A58.25	G1	F	No	Yes
PLC A58.28	G1	F	Yes	No
PLC A58.29	G1	F	No	Yes

Table 5.1. Segregation of multiple arrays. The PLC A58 founder contained two 'reduced arrays' which when this animal was bred to the next generation, some animals inherited the 10kb RV locus and some inherited the 17kb RV locus. No G1 animal inherited both. 18 out of the 29 G1 tested were transgenic (transmission of 62%).

The animals were also screened for the presence of the PGK-Cre construct but no animals were found to have an integrated copy of the PGK-Cre construct (data not shown). Therefore, the strategy of microinjection of a circular plasmid to avoid integration of the construct, appears to have been successful in these animals.

Animals from the lox 5 line microinjected with Cre were termed PLC B animals. Only one animal from lox 5/PGK Cre (PLC B) showed reduced arrays, PLC B31. The array was estimated to be approximately 17kb, which is equivalent to the larger arrays found in the PLC A58 and A39 animals (figure 5.1). Other animals appeared to have smears of fragments that may indicate that the Cre recombination was producing different sized arrays in different cells throughout the animal. The reduced animal was used to establish a line but this has presently been unsuccessful, as only one positive G1 animal out of 18 screened has inherited the reduced array (figure 5.2).

The inability to produce large numbers of animals with 'reduced arrays' in the microinjected lox 5 animals (from 43 positive animals that had been microinjected only one, PLC B31, showed a 'reduced array') is interesting in the context of the results from Chapter 4. When lox 5 animals were bred to a BLG-Cre line there was no evidence of a 'reduced array' in any animal. However, there had been a reduction in the intensity of the 'parental array' suggesting that reduction in copy number had occurred. The inability of lox 5 animals to generate 'clean' reduced arrays after Cre expression, suggests that there is something inherent to the original structure of the lox 5 array which may have produced unpredicted Cre mediated recombination.

Due to the inability to establish a line of reduced animals from the lox5/PGK Cre founders, PLC B animals will not be discussed further.

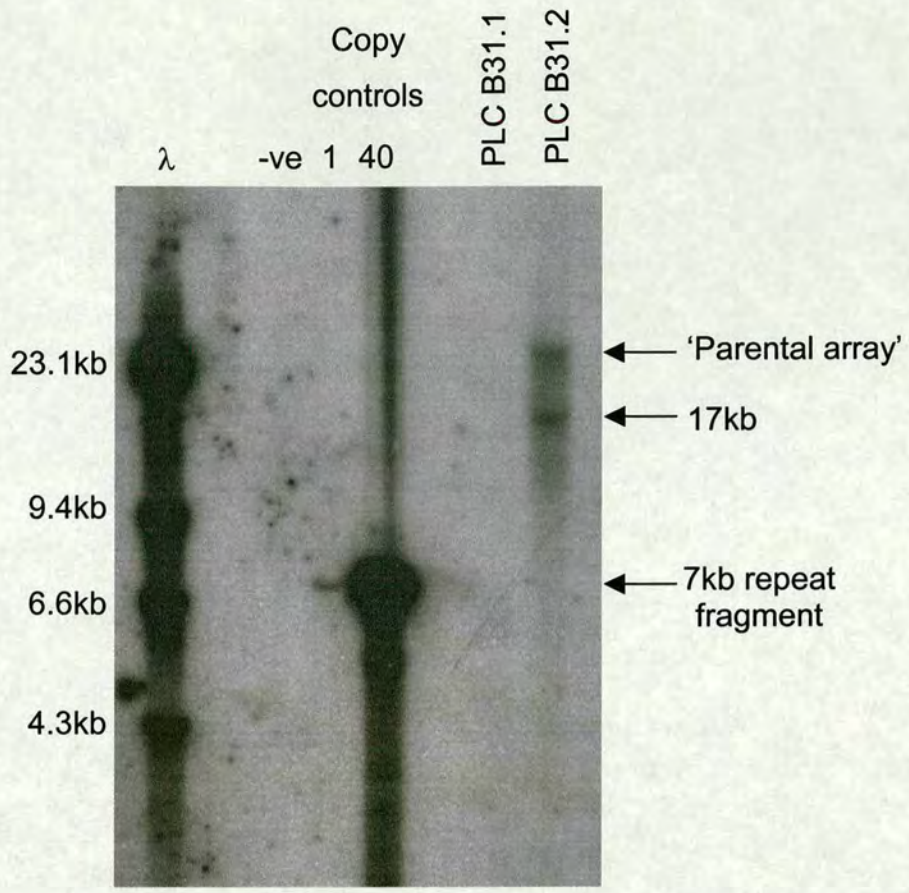


Figure 5.2. EcoRV digestion of PLC B G1 mice. 10μg tail DNA from PLC B mice was digested with EcoRV, electrophoresed on a 0.7% gel, blotted to Zetaprobe and hybridised to a 1.1kb fragment of BLG (described in Chapter 2). Copy controls were dilutions of the original microinjected 7kb BLG-loxP fragment added to 10μg of non-transgenic DNA. –ve indicates non-transgenic DNA. λ indicates a lambda HindIII ladder with stated sizes. PLC B31.1 was a non-transgenic sib to the only reduced G1 PLC B animal, PLC B31.2.

Similar to the majority of the lox 5/PGK Cre animals, lox 9/PGK (PLC C) animals showed no clear reduced array (figure 5.3). These animals showed complex smears of fragments. The smear may be due to different recombinations occurring in different cells. This may also reflect continued expression of Cre recombinase over a period of development allowing some cells to undergo several rounds of recombination. The use of a 0.7% gel may

have reduced the ability to resolve fragments of similar size. This 'smear' could theoretically represent non-specific degradation of the DNA by EcoRV. However, the samples of figure 5.1 were digested with the same batch of enzyme under the same conditions, making this seem unlikely. As with lox 5 animals, in both types of Cre recombination (in the mammary or the embryo) lox 9 animals failed to produce a 'reduced array'. The structure of this line may mean that recombination to a single copy array is not possible. Due to the failure to find any animal with a 'reduced array', these animals were not further investigated.

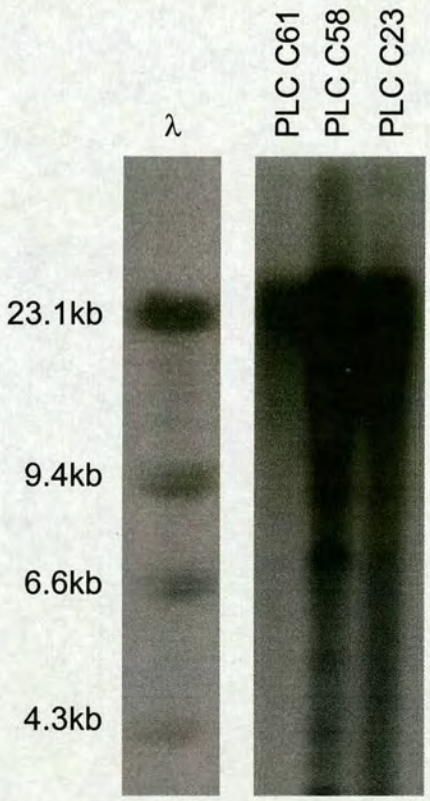


Figure 5.3. EcoRV digestion of microinjected PLC C founder mice. 10μg tail DNA from PLC C founders was digested with EcoRV, electrophoresed on a 0.7% gel, blotted to Zetaprobe and hybridised to a 1.1kb fragment of BLG. λ indicates a lambda HindIII ladder with stated sizes.

Southern blotting with 5', middle and 3' probes was carried out to determine whether the 'reduced array' fragments contained an intact BLG-loxP transgene. If the 'reduced arrays' carried an intact transgene it would be predicted that all three probes would hybridise to the fragment. From figure 5.4, it is clear that the 10kb RV locus and the 17kb RV locus both hybridise to all three probes. This indicates that both the 'reduced arrays' contained the sequences corresponding to the transgene probes. Although the 10kb locus animals were derived from different founders, because they showed the same EcoRV pattern they were assumed to contain the same 'reduced array'. All the 17kb RV locus animals were derived from the same founder.

Putative maps of the structure of the 'reduced arrays' are shown in figure 5.5. However, this data could not formally exclude the possibility that there had been rearrangements or mutations in the transgene.

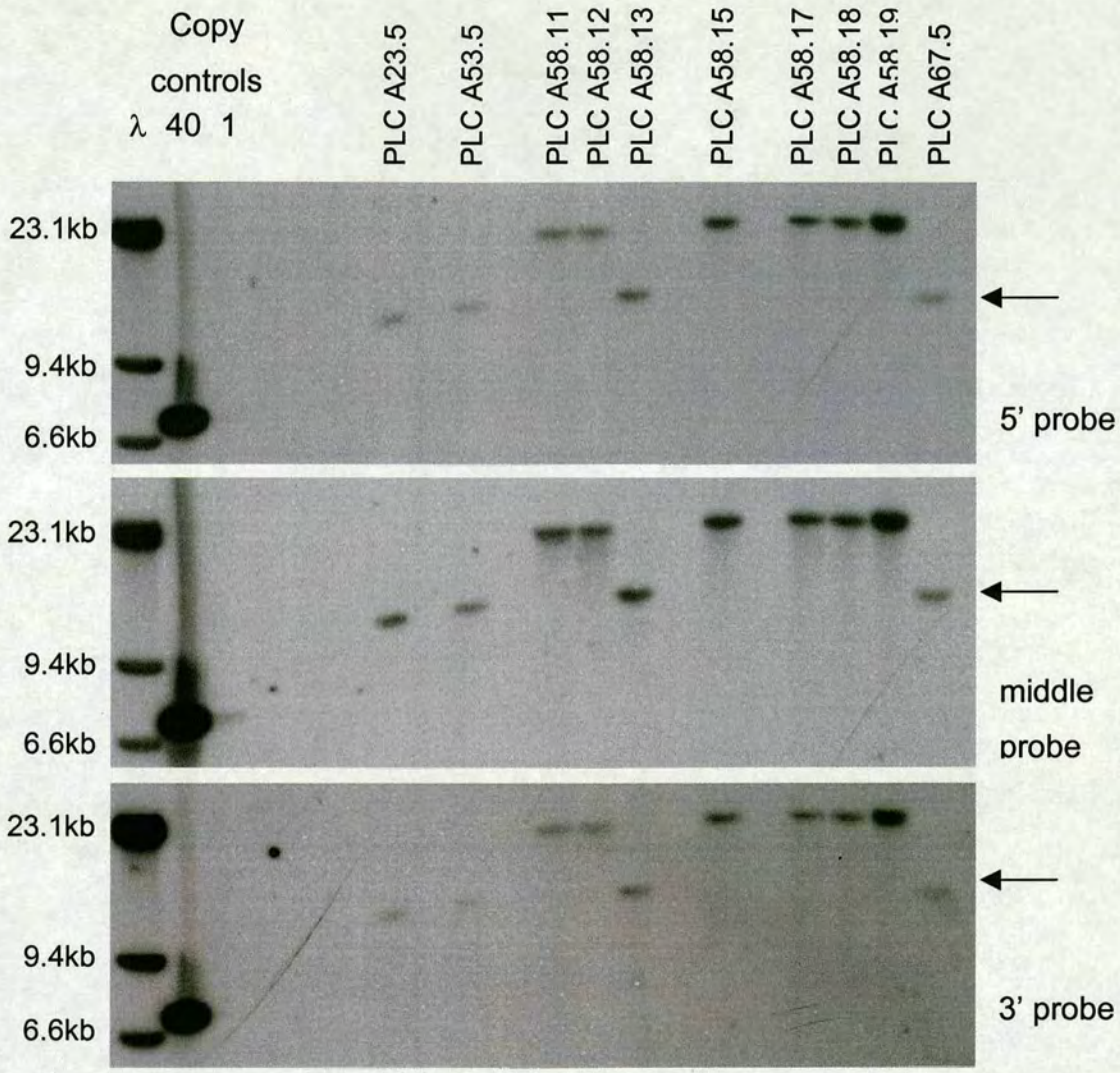


Figure 5.4. Multiple hybridisations of EcoRV digested PLC A DNA. 10 μg of tail DNA was digested with EcoRV, electrophoresed on a 0.7% gel, blotted to Zetaprobe and hybridised sequentially with 5', middle and 3' probes to BLG. λ indicates a lambda HindIII ladder with stated sizes on left hand side. The 10kb RV locus is indicated with a black arrow. Lanes without labels are negative sibs.

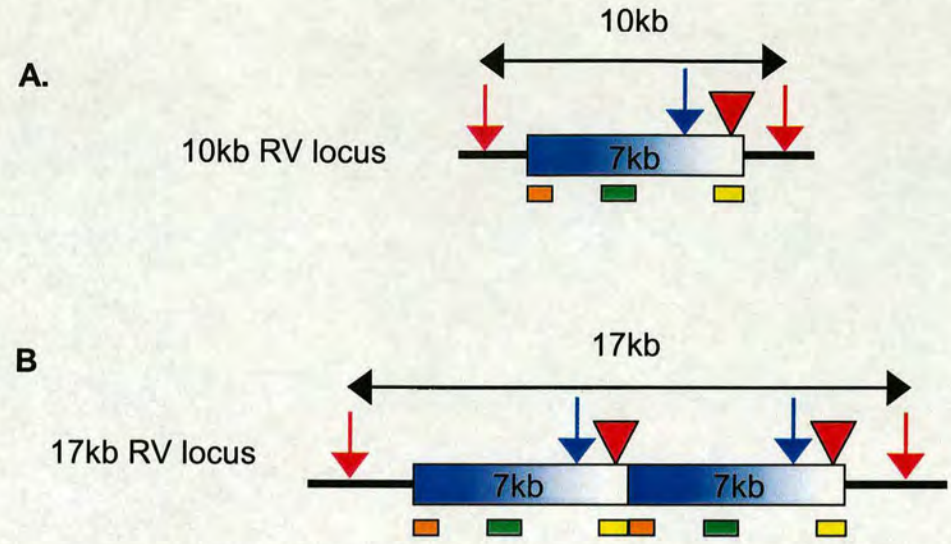


Figure 5.5. Putative maps of EcoRV Southern blot for 10kb RV and 17 RV locus. If either the 10kb RV locus (a) or 17kb RV locus (b) contain a whole transgene they would be expected to hybridise to all three probes (5' probe: orange box, middle probe: green box, 3' probe: yellow box. Probes detailed in figure 3.5 and Chapter 2). The red triangle represents the loxP site. The red arrows represent the EcoRV sites within the genome, and the blue arrows represent the SspI site within the BLG transgene (discussed in reference to figure 5.7). The sizes given for the arrays are those determined from figure 5.1.

Further Southern blotting was carried out to determine more about the structure of the 'reduced arrays'. SspI cuts only once in the BLG-loxP transgene and had been used in both Chapter 3 and 4 to estimate the copy numbers of the transgenic lines. If there are two copies of the transgene present in a head to tail repeat, SspI digestion should produce a 7kb fragment and the transgene junction fragment. If the fully 'reduced arrays' only contained one transgene no repeat length fragment would be produced (figure 5.6). This was used to determine the structure of the BLG transgene in the 10kb RV locus and the 17kb RV locus (figure 5.7).

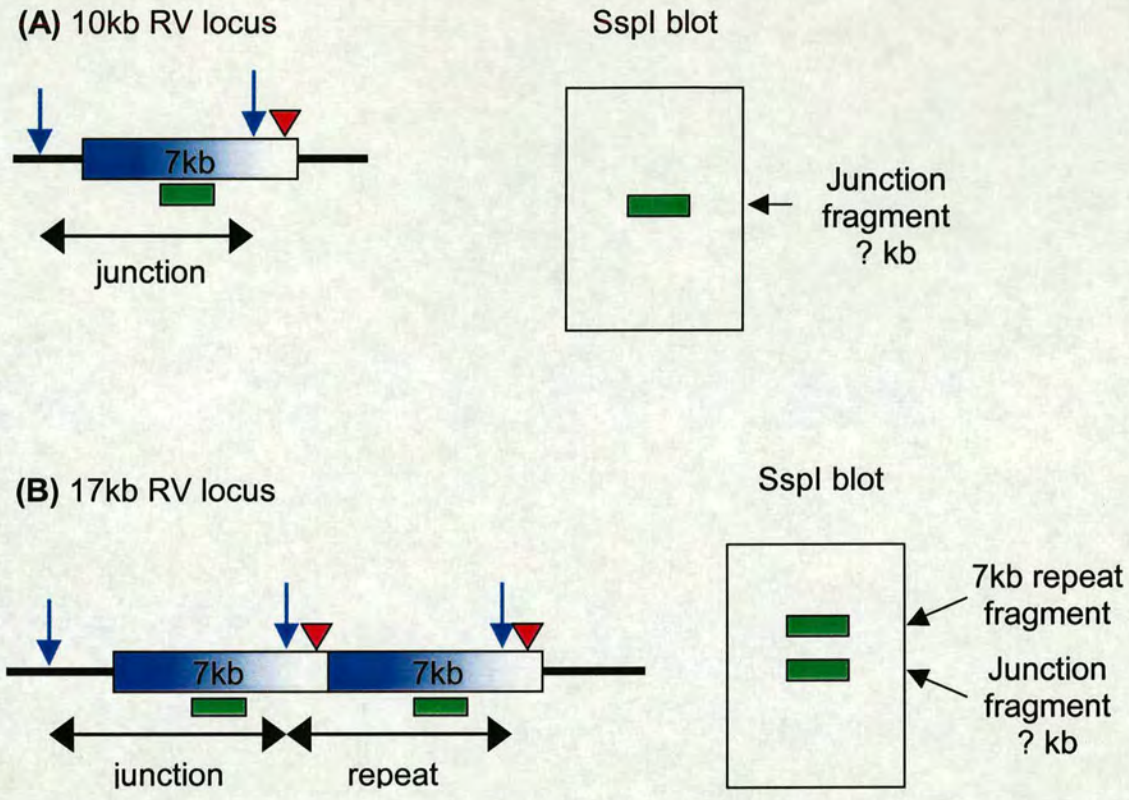


Figure 5.6. Diagram of production of a repeat length fragment from a transgenic locus (A) If there is only one transgene within the reduced array it cannot produce a repeat fragment. (B) If there are two or more transgenes within a 'reduced array' in a head to tail manner it will produce a 7kb repeat fragment after SspI digestion. The black line represents genomic sequence, the blue arrow represents the SspI restriction site (there is only one within the BLG-loxP transgene), the green box represents the probe, the red triangle represents loxP site.

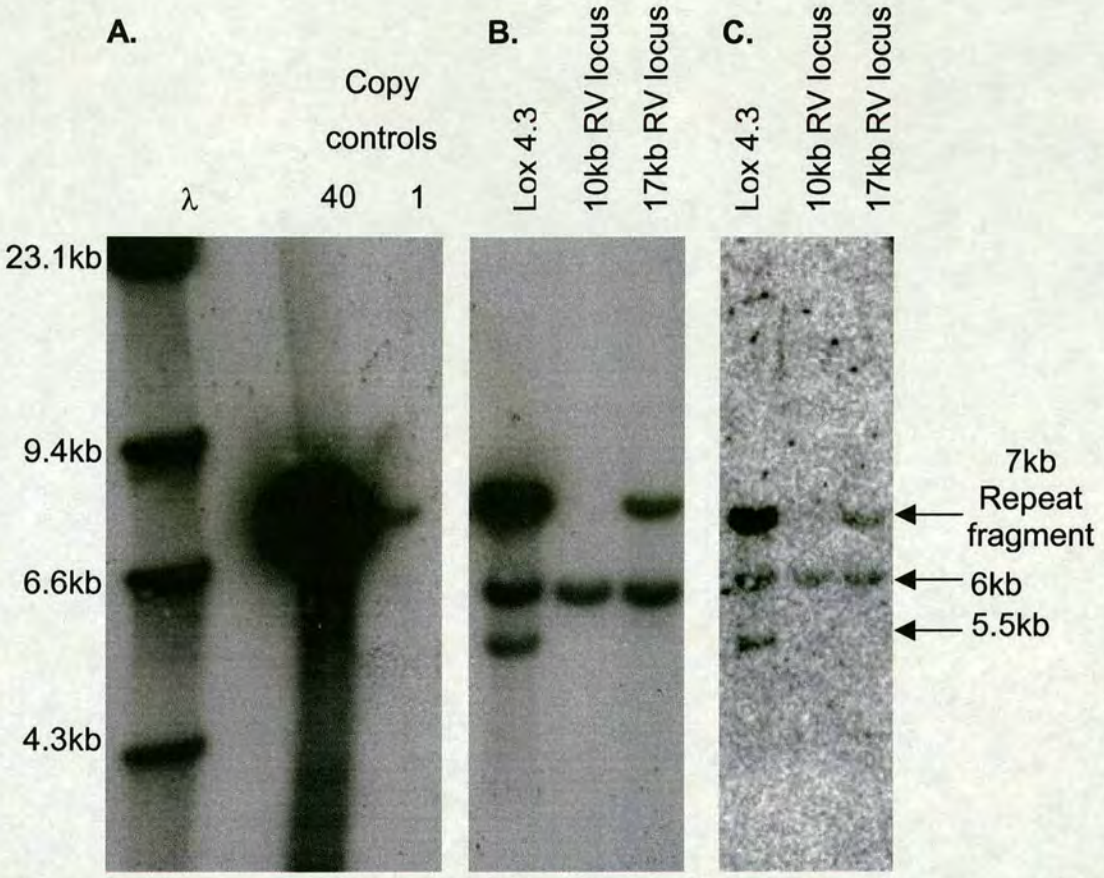


Figure 5.7. SspI digestion of PLC A DNA to determine the copy number of the reduced array. 10μg of liver DNA was digested with SspI, electrophoresed on a 0.7% gel, blotted to Zetaprobe. Copy controls were dilutions of the original microinjected 7kb BLG-loxP fragment added to 10μg of non-transgenic DNA (panel A, hybridised with the middle probe). The transgenic DNA was hybridised to the middle probe (panel B) or the 5' probe (panel C) described in figure 3.5 and Chapter 2. The blot was stripped between probes and exposed to X-film to confirm the removal of the original probe. λ indicates a lambda HindIII ladder with stated sizes. The 6kb is a presumed junction fragments. The 5.5kb fragment is thought to be an internal truncated fragment as discussed in figure 5.9.

From figure 5.7 the SspI digestion shows that the 10kb RV locus does not contain a repeat length fragment. This also correlates with its EcoRV size (figure 5.1 and 5.5), if there were two transgenes within the EcoRV fragment it would be predicted to be above 14kb (the transgene is 7kb long). This

confirms that there are less than two intact copies of the transgene present in the array, as the repeat length fragment is only present when at least two input transgenes integrate in a head to tail array. The 6kb SspI fragment hybridises to both the 5′ probe and the middle probes. The SspI site cuts 5.4kb from the 5′ end of the BLG transgene. Therefore, the fact that the SspI fragment is longer than 5.4kb suggests that it contains the whole transgene fragment up to the SspI site as well as some genomic sequence. This is shown in figure 5.8.

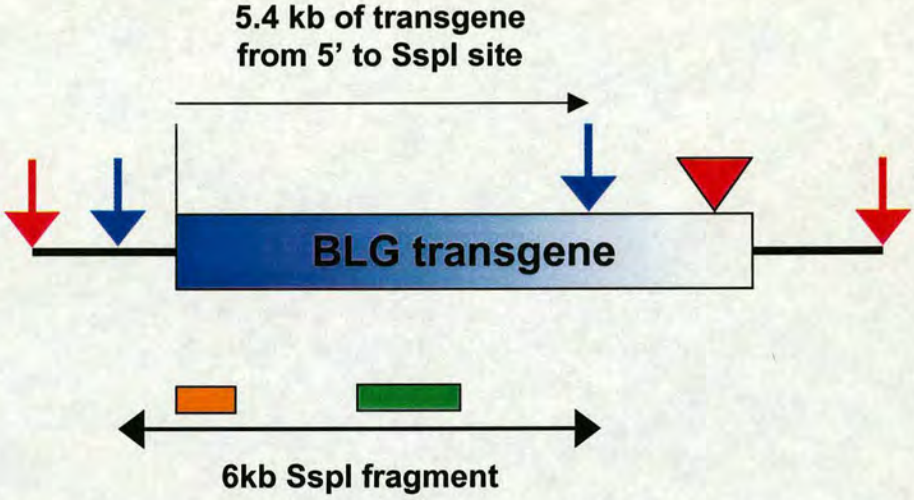


Figure 5.8. Putative model to account for the 6kb SspI fragment from 10kb RV locus animals. The reduced transgenic locus was originally defined by the size of the EcoRV fragment (genomic EcoRV sites are represented by red arrows, the lox P site is represented by red triangle). Within this, there is a BLG transgene which contains only one SspI site 5.4kb from the extreme 5' end of the transgene (blue arrow). SspI digestion produces a 6kb fragment from the 10kb RV locus (double headed black arrow). This 6kb fragment hybridises to both the 5' probe and the middle probe (orange and green boxes, probes described in figure 3.5 and Chapter 2). As the SspI fragment is longer than 5.4kb this suggests that the transgene is intact from the 5' to the SspI site.

Animals containing the 17kb RV locus (i.e. PLC A58.8) showed that they had a 7kb repeat length fragment from SspI digestion (figure 5.7). The presence of one intact transgene is confirmed by the presence of a repeat length fragment. This indicates that there may be at least two transgenes present in a head to tail array. Comparisons between the intensity of the repeat fragment and the 6kb fragment shows that they are of comparable intensity. If the 6kb fragment represents a junction fragment, given that junction fragments are presumed to exist only as a single copy per genome, this suggests that there is only one repeat unit present in the array.

In figure 5.7, the 5.5kb fragment is also lost after the Cre reduction, which may indicate that within the array there exists a truncated copy of the transgene that is removed upon Cre recombination. A model of a possible structure to account for this is shown in figure 5.9.

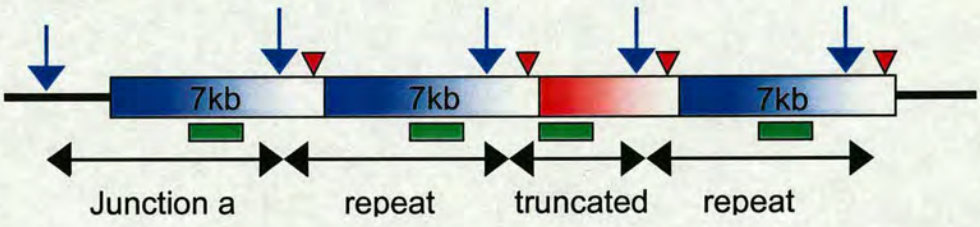


Figure 5.9. Possible structure of lox 4 transgenic loci. If there was a truncated transgene (red box) that still contained the SspI site (blue arrow) and a loxP site (red triangle) then after Cre mediated recombination it would be completely removed. This would leave junction a, and a repeat length fragment if the array contained more than one complete 7kb transgene. The green box represents the middle probe used for figure 5.7.

5.2.2 Long Range PCR

Although the data from figure 5.7 suggested that there was an intact transgene present within the 'reduced array', more work was carried out to confirm the integrity of the transgenes within the 10kb RV locus and the 17kb RV locus. Long range PCR was used to determine if there was a continuous 7kb intact BLG-loxP transgene present within the reduced arrays. If there was a PCR product this method could also provide the possibility of cloning the product allowing it to be analysed further. The cloned PCR product could be sequenced to determine if it was rearranged or mutated. Alternatively it could be used in transient transfections of mammary cells *in vitro* to formally prove if it was still capable of expressing after being reduced by Cre recombinase.

Primers were designed for the extreme ends of the 7kb BLG-loxP injection fragment (figure 5.10) and long range PCR was carried out on the 10kb RV locus and the 17kb RV locus. PCR was also carried out on the original microinjection 7kb fragment, lox 4 parental DNA, and double transgenic lox4/Cre DNA. Distilled water and negative DNA were used to control against contamination. If the array contained an intact BLG-loxP transgene, a 7kb fragment would be predicted as the product.

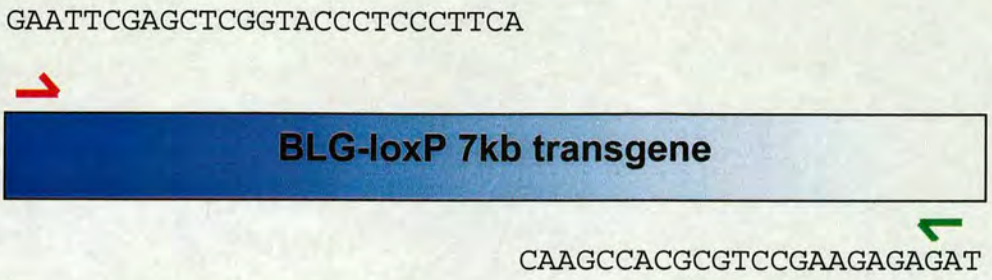


Figure 5.10. Primers used for long range PCR of reduced array DNA. The red arrow represents the 5' junction primer, which was designed to prime from 12bp from the 5' end of the transgene sequence. The green arrow represents the 3' primer which was designed to prime 79bp from the 3' end of the transgene. The sequence of the primers is given beside the arrows.

A 7kb fragment was derived from the 17kb RV locus, but no visible product was produced by the 10kb RV locus. The PCR products were also blotted to determine if there was low-level product but no fragment was detected from the 10kb RV locus (figure 5.11).

The failure to successfully PCR a 7kb fragment from the 10kb RV locus suggests that this 'reduced array' does not contain an intact BLG-loxP transgene. Interestingly no 7kb product was found from lox 4/Cre double transgenic DNA (GB6.7; figure 5.11). This may indicate that the 10kb 'reduced array' in Chapter 4 also did not contain an intact transgene. Although the Southern blotting with the 10kb RV locus was consistent with the 'reduced array' containing all the BLG transgene sequence it is possible that the 'reduced array' does not contain an intact transgene. However, the 17kb RV locus does appear to contain an intact BLG-loxP transgene, not only by the PCR data but also by the finding that it contains a 7kb SspI repeat fragment.

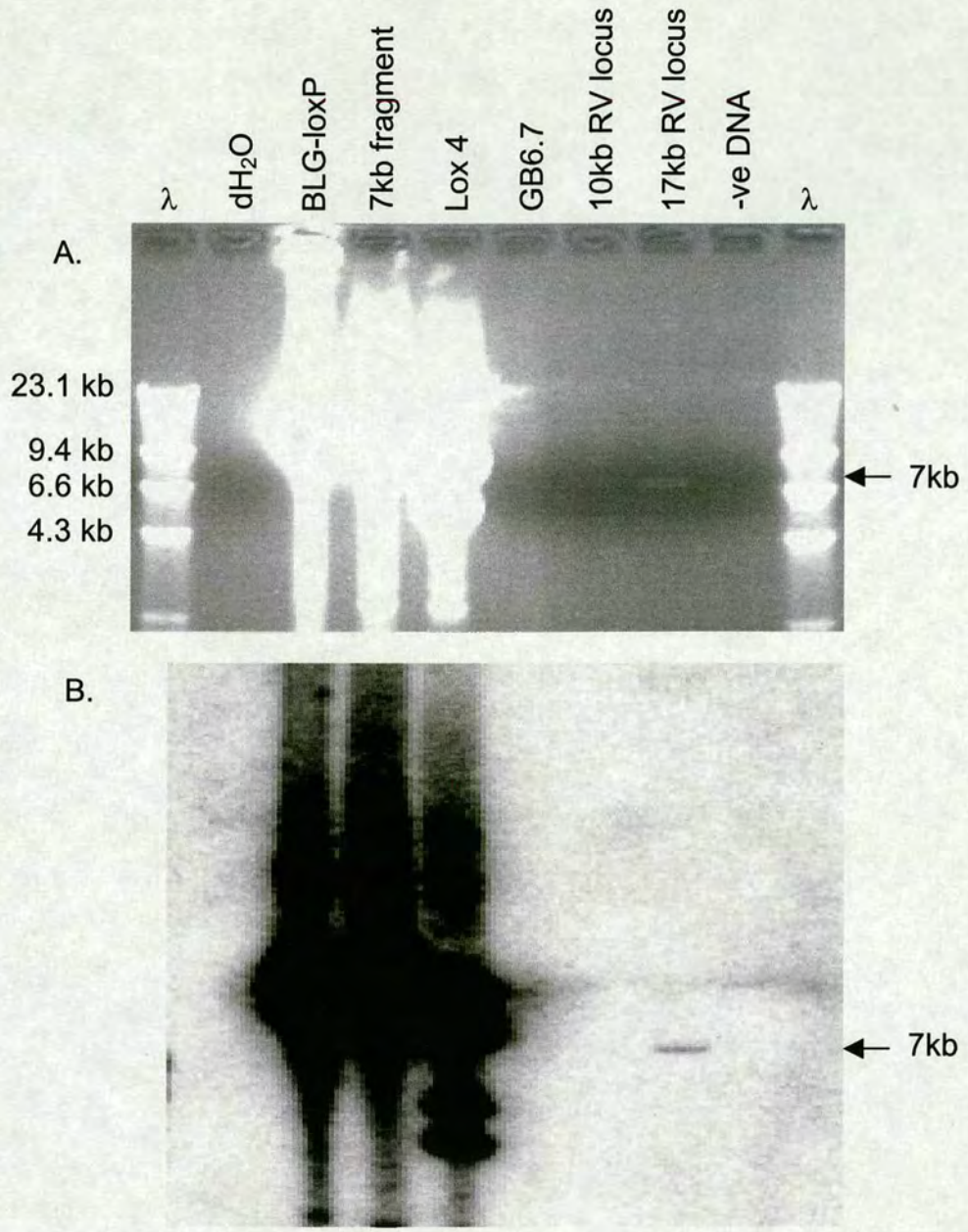


Figure 5.11. Long Range PCR on reduced transgenic DNA. (a) Ethidium bromide staining of long range PCR products. (b) Southern blot of the same gel showing hybridisation of a 1.1kb fragment to the 7kb PCR product in 17kb RV locus DNA. λ indicates a lamda HindIII ladder with stated sizes. Distilled water and negative DNA are used as a control against PCR contamination. BLG-loxP refers to the plasmid used to derive the BLG-loxP transgene. The 7kb fragment was the original microinjected fragment. Mammary DNA was used from lox 4, a lox4/Cre double transgenic (GB6.7) and either 10kb RV or 17kb RV locus animals.

5.2.3 Expression Analysis

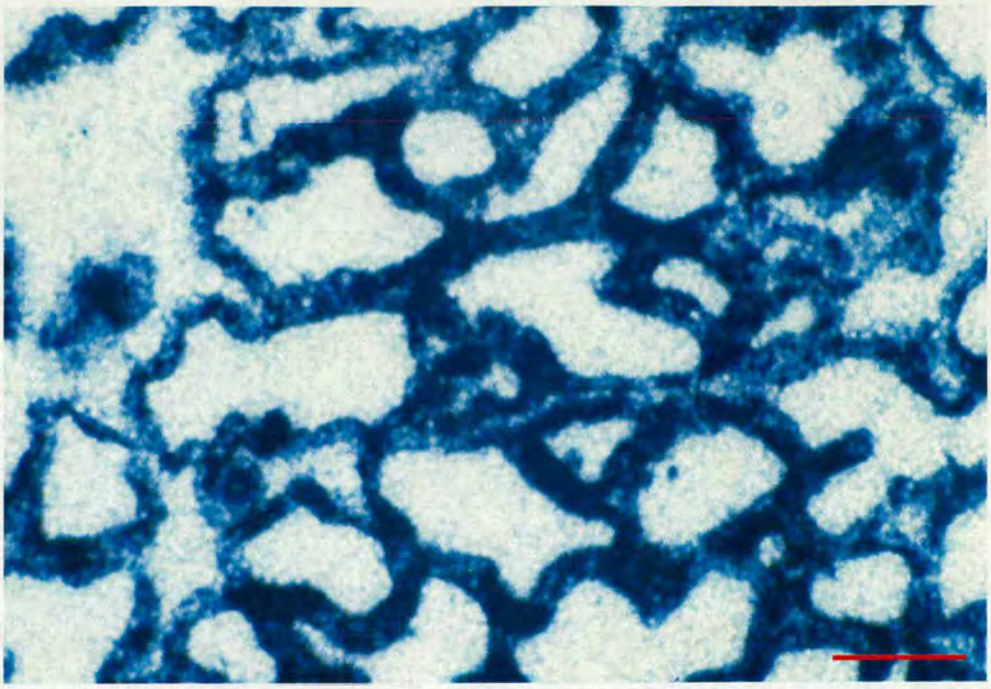
5.2.3.1 *In situ* hybridisation

To determine the cellular pattern of the transgene expression after the reduction of copy number by microinjection of Cre recombinase, *in situ* hybridisation was carried out as described in 2.8.

Controls against background (sense RNA probes) and negative mammary section were always negative for BLG expression even when photographed using dark field microscopy (figures 5.14 and 5.17), indicating as in other chapters there was no non-specific hybridisation.

Sections from 10kb RV locus animals showed only occasional spots of BLG expression (figure 5.12). These spots covered only single cells and were found very infrequently. Figure 5.12 represents the strongest collection of dots found in any section of the 10kb RV locus animals analysed. During *in situ* hybridisation sequential sections were placed on a slide so that for each probe there were two sections per slide. In previous experiments with the lox lines there was always a consistent pattern between the sequential sections. This was not the case in these animals, the patterns of dots was not mirrored between sections, indicating that it may be an artefact of the *in situ* technique rather than genuine expression of the transgenic mRNA. Figure 5.13 shows that in another area of section no BLG expression could be found by dark field microscopy.

A.



B.

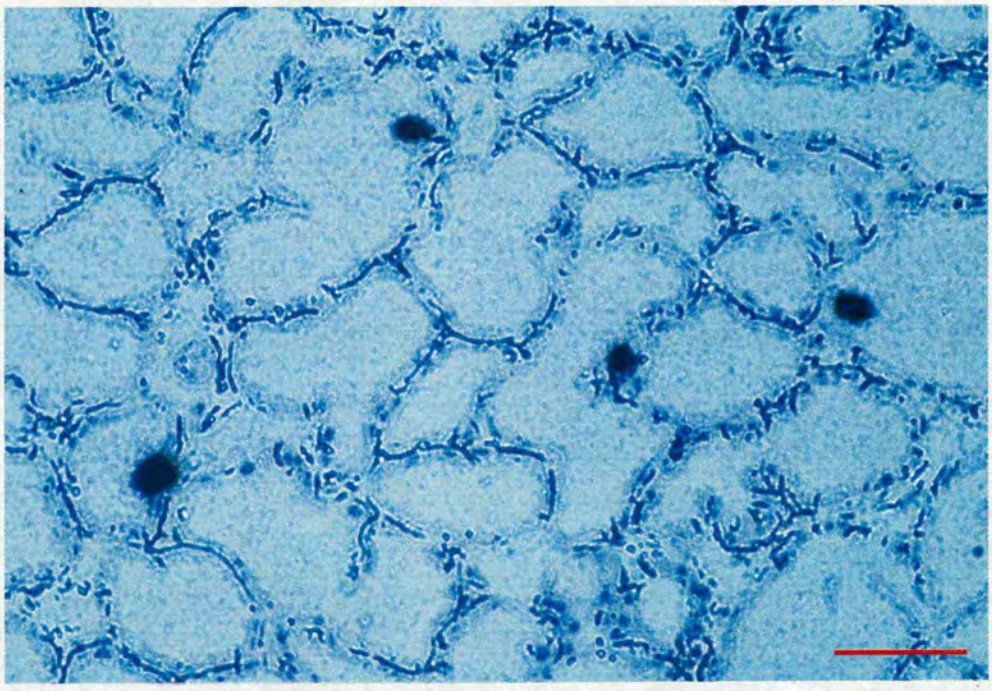


Figure 5.12. *In situ* hybridisation analysis of mRNA expression patterns in a 10kb RV locus animal. (A) β-casein, (B) BLG. Photographs were taken after 6 week exposures using a x20 objective. Bar represents 150μm.

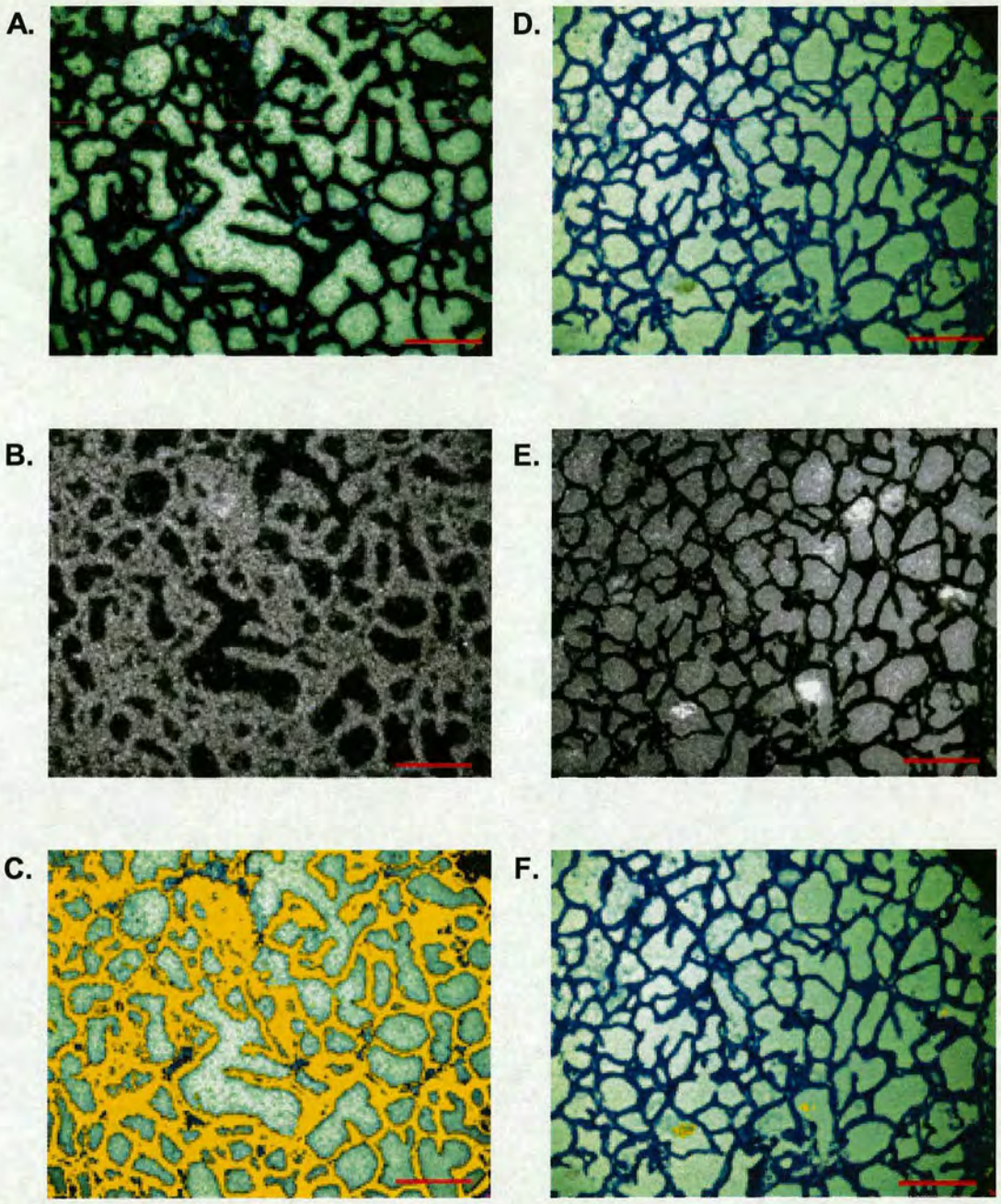


Figure 5.13. Dark field imaging of *in situ* hybridisation of 10kb RV locus tissue. (A,B,C) probed with β-casein. (D,E,F) probed with BLG. Photographs were taken after 2 weeks exposures using a x10 objective. (A) and (D) are light field images. (B) and (D) are dark field images. (C) and (F) are composite images of both light and dark field with the silver grains being represented with a yellow paint. Bar represents 300μm.

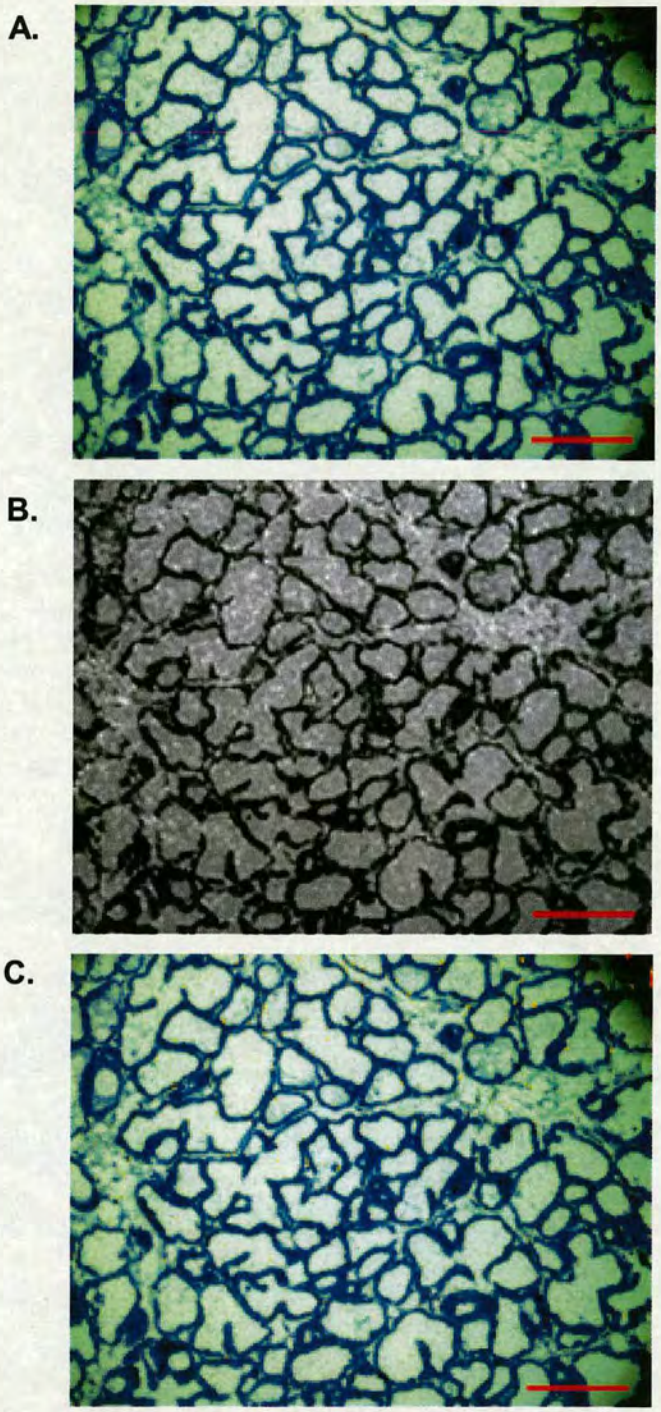
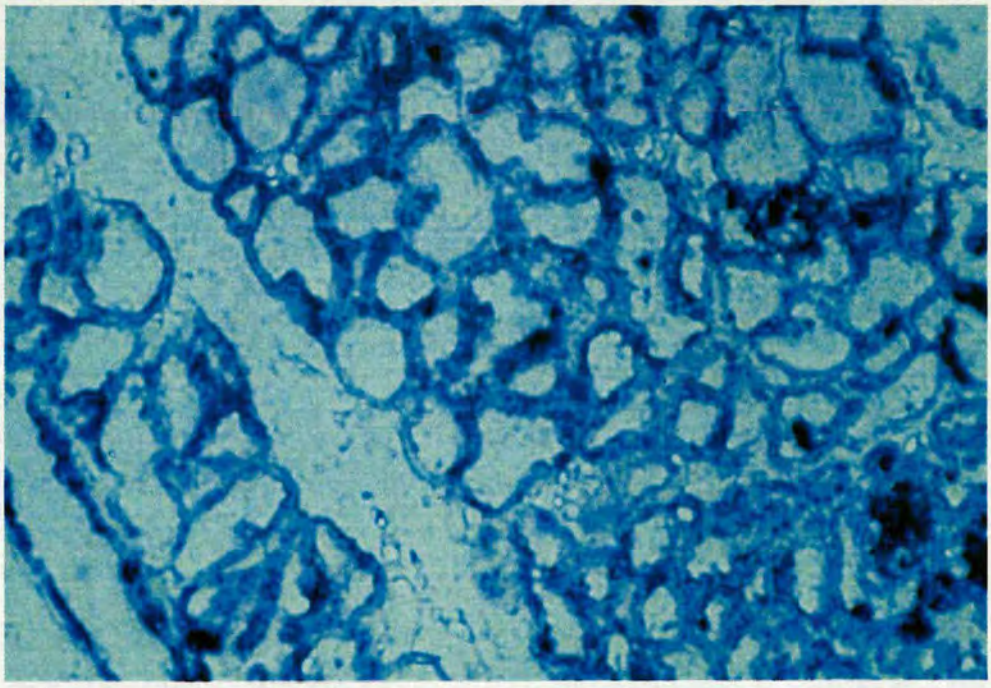


Figure 5.14. Control *in situ* hybridisation of 10kb RV locus tissue. (A,B,C) probed with sense probe. Photographs were taken after 6 weeks exposures using a x10 objective. (A) is a light field image. (B) is a dark field image. (C) is a composite image of both light and dark field with the silver grains being represented with a yellow paint. Bar represents 300μm.

A.



B.

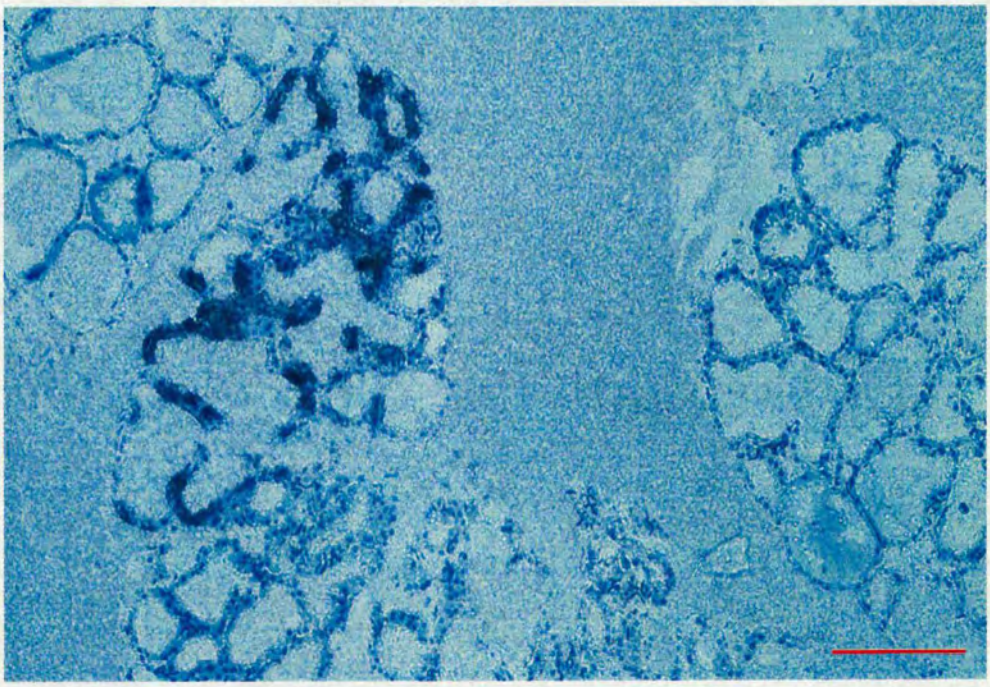


Figure 5.15. *In situ* hybridisation analysis of mRNA expression patterns in a 17kb RV locus animal. (A) β-casein, (B) BLG. Photographs were taken after 6 week exposures using a x10 objective. Bar represents 300μm.

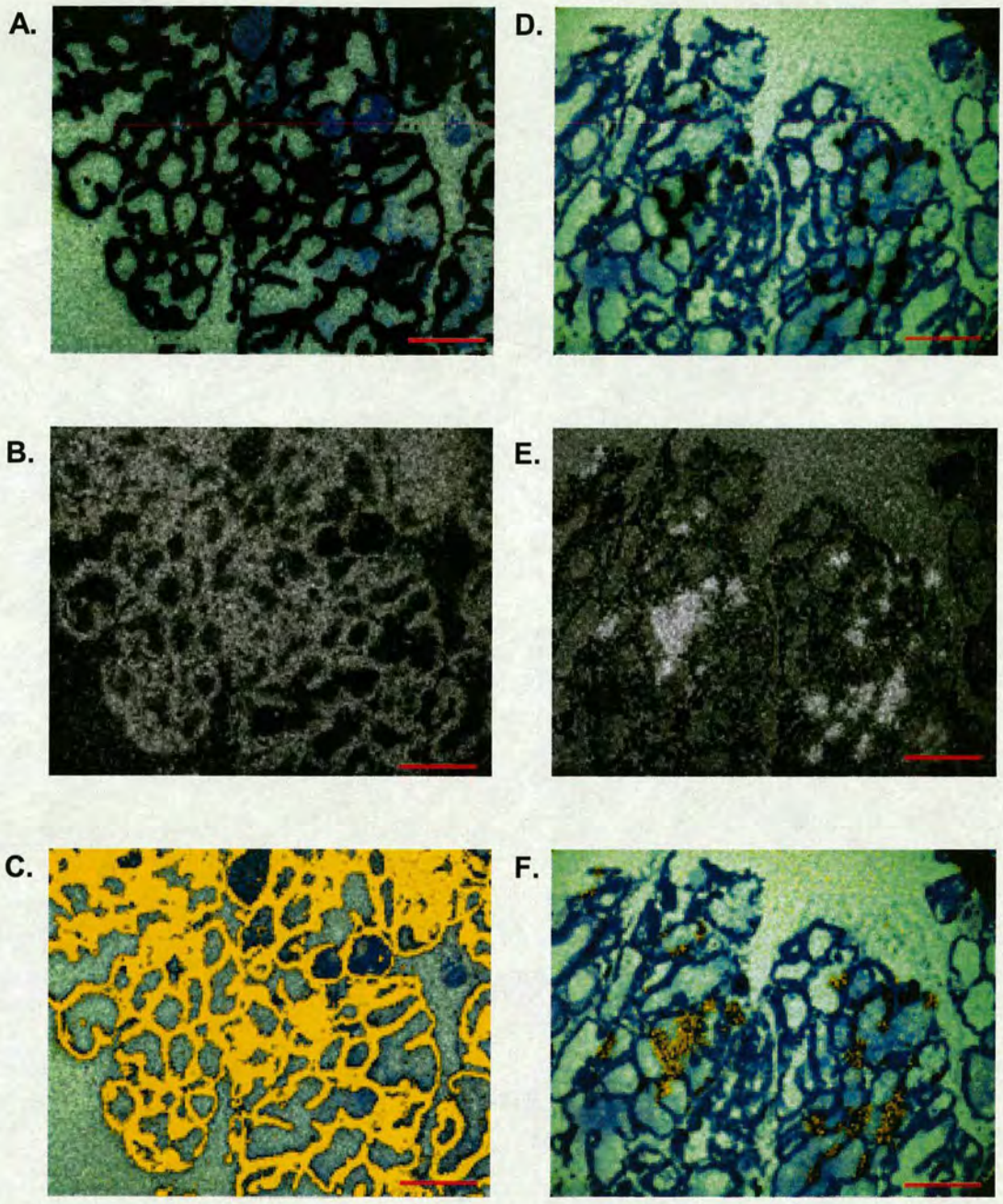


Figure 5.16. Dark field imaging of *in situ* hybridisation of 17kb RV locus tissue. (A,B,C) probed with β-casein. (D,E,F) probed with BLG. Photographs were taken after 2 weeks exposures using a x10 objective. (A) and (D) are light field images. (B) and (D) are dark field images. (C) and (F) are composite images of both light and dark field with the silver grains being represented with a yellow paint. Bar represents 300μm.

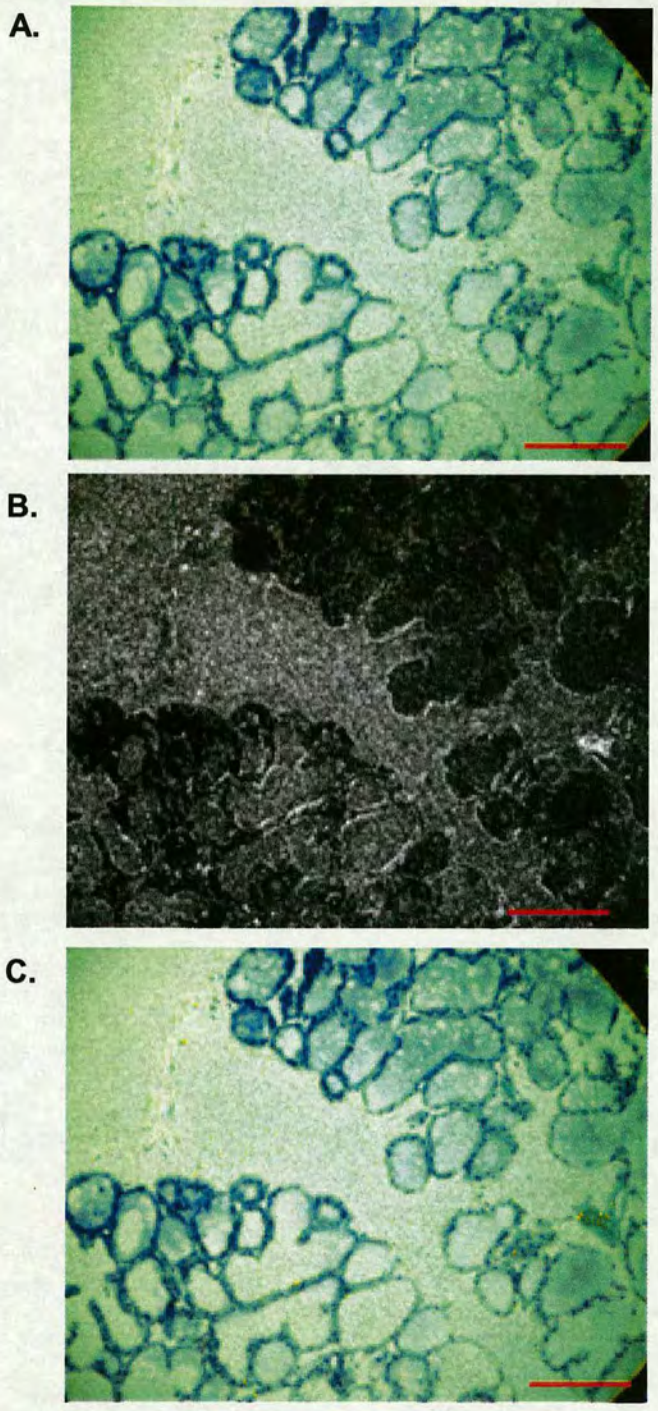


Figure 5.17. Control *in situ* hybridisation of 17kb RV locus tissue. (A,B,C) probed with sense probe. Photographs were taken after 6 weeks exposures using a x10 objective. (A) is a light field image. (B) is a dark field image. (C) is a composite image of both light and dark field with the silver grains being represented with a yellow paint. Bar represents 300μm.

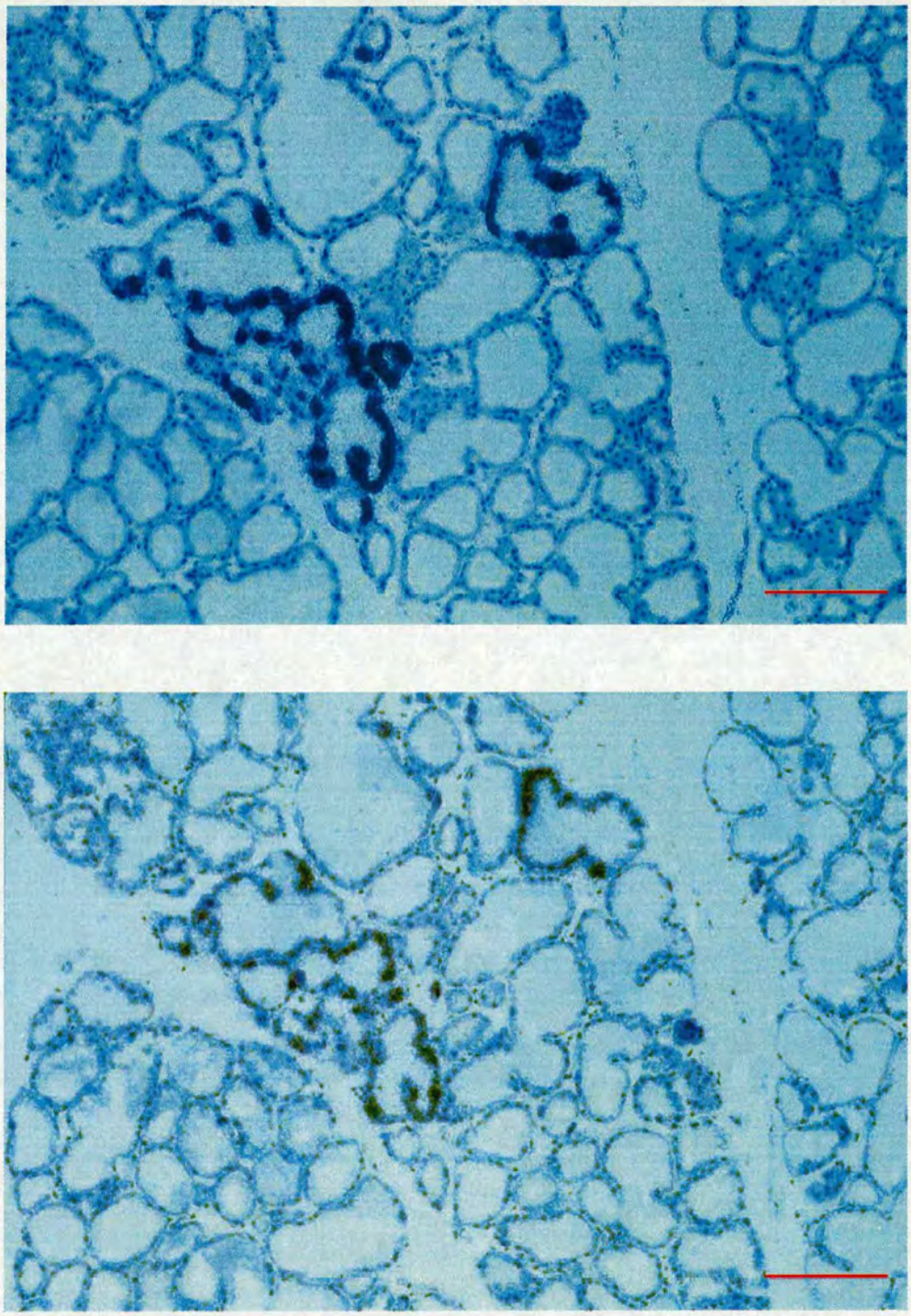


Figure 5.18. *In situ* hybridisation of sequential sections in a 17kb RV locus animal. Both sections were probed with BLG. Photographs were taken after 2 week exposures using a x10 objective. Bar represents 300μm.

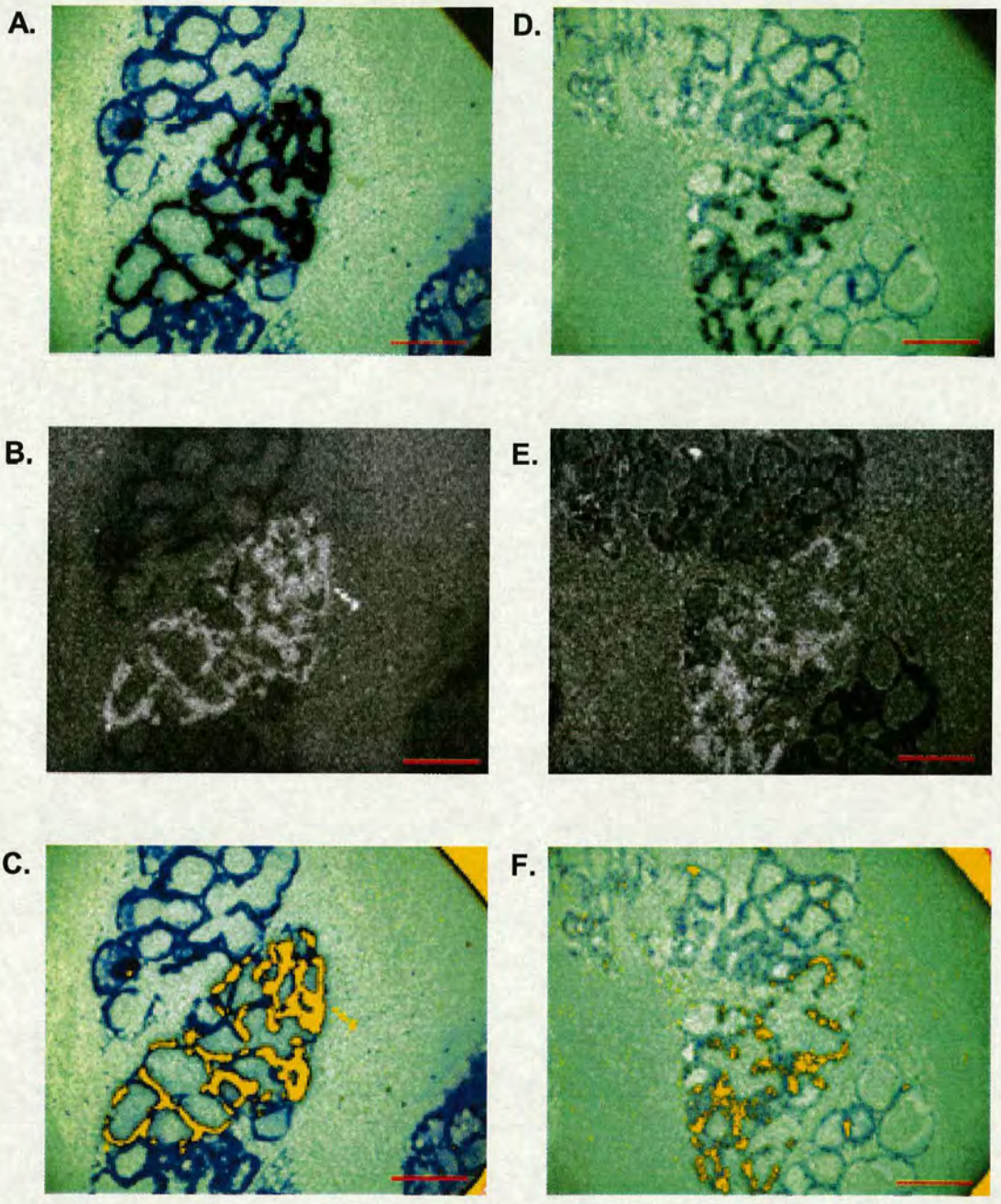


Figure 5.19. Dark field imaging of *in situ* hybridisation of sequential sections from a 17kb RV locus animal. (A,B,C) probed with BLG. (D,E,F) probed with BLG. Photographs were taken after 2 weeks exposures using a x10 objective. (A) and (D) are light field images. (B) and (D) are dark field images. (C) and (F) are composite images of both light and dark field with the silver grains being represented with a yellow paint. Bar represents 300μm.

By contrast, sections from 17kb RV locus animals showed patches of expression (figure 5.15). These patches of expression were also visible by dark field microscopy (figure 5.16). The pattern was consistent across sequential sections indicating that the expression was genuine (figures 5.18 and 5.19). These patches of expression were found in all of the 17kb RV locus animals examined (n=6). The dark spotting in figure 5.15 under light field may be due to clumps of cells that have become more heavily counterstained. This effect is not seen under dark field (figure 5.16).

Like the results discussed in Chapter 4 this result contradicts published work (Garrick *et al* 1998). The finding that the 10kb RV locus are silent while 17kb RV locus shows only limited patchy expression may suggest that there is a requirement for more than one BLG transgene at this genomic location to allow expression. The alternative explanation for the lack of expression from the 10kb RV locus is because it is functionally inactivated. However, the 17kb RV locus by Southern blotting and PCR analysis appears to show an intact transgene and it is still expressing at a much lower level than the parental lox 4 line. It is clear that this BLG transgene at this genomic location expresses more efficiently as a multicopy array rather than a two copy array.

As with the *in situ* analysis in Chapter 3 and 4, the amount of cells expressing BLG was quantified. Alveoli were scored according to the amount of area seen to be expressing BLG RNA; negative when no cells in the alveolus expressed, -/+ when there was partial expression from an alveolus, and positive when all the cells in an alveolus expressed. Figures giving the percentages of alveoli in each group are given in table 5.2. The rare spots

from the 10kb RV locus animals was taken to represent background and these animals were not used for quantitative analysis.

The counts were analysed using GenStat software to produce Table 5.3 combining the values of alveoli counted as partly positive and those fully positive. Standard deviations (s.d.) and coefficient of variance are both measures of the amount of variability within a group i.e. the differences between individuals from the same transgenic line.

Animal ID	*Status*	*Negative*	*Partly*	*Positive*
PLCA58.11	17kb RV locus	93.2	5.4	1.4
PLCA58.12	17kb RV locus	93.7	5.3	1.0
PLCA58.15	17kb RV locus	95.5	4.1	0.4
PLCA58.17	17kb RV locus	99.3	0.6	0.1
PLCA58.18	17kb RV locus	95.4	4.6	0.0
PLCA58.19	17kb RV locus	98.5	1.4	0.1
PLCA58.25	17kb RV locus	92.0	6.9	1.1

Table 5.2. BLG expression values by *in situ* hybridisation of 17kb RV locus animals. Ten fields per section and two sections for each animal were counted to derive a percentage for the three categories. The status refers to the type of 'reduced array' carried by the animal.

Status	*Range of percentages of positive alveoli*	*n*	*Mean*	*s.d*	*c.v*
17kb RV locus	0.7-8.0	7	4.6	2.7	58.7

Table 5.3. Statistical analysis of 17kb RV locus transgenic *in situ* counts. Range is given as a percentage of the total alveoli which had some expression, *n* = number of animals per line, *s.d.* standard deviation, *c.v.* = 100 x s.d. / mean.

The range of total positive alveoli (partial or complete) for the 17kb RV locus animals is shown graphically in figure 5.20. It also compares the 'reduced array' animals to the parental lox 4 line. There is a significant decrease from the parental lox 4 line which expressed in approximately 50% of the alveoli, to the 17kb RV locus animals which all express at less than 10% of the alveoli. The 10kb RV locus animals only expressed at less than 1% of the alveoli counted, which is more consistent with background rather than genuine expression.

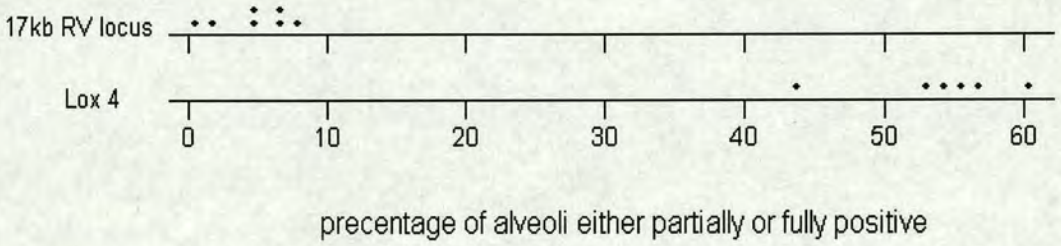


Figure 5.20. Dotplot comparing the 17kb RV locus reduced animals with the parental lox line. Each dot represents an animal. The percentage alveoli that were counted partial or fully positive were combined to give an overall view of the total expression from the mammary gland.

The microinjected lines were statistically compared with the parental line lox 4 to see if there were differences. There is a strong statistical difference (p=0.003) between the parental lox 4 line and the 17kb RV locus animals in the percentage of alveoli that express the transgene.

5.2.3.2 Northern blots analysis

Northern blots were carried out to determine whether or not the levels of variegation shown by *in situ* analysis was mirrored in total RNA isolated from mid-lactation mammary tissue. As with the case of the double transgenic animals in Chapter 4, no BLG RNA could be detected by this method.

5.2.3.3 Milk protein analysis

SDS-PAGE analysis was used to determine if these reduced animals expressed BLG at the protein level. Mouse milk samples were prepared as described (see 2.7.1) and were loaded at 1/250 or 1/100 dilutions along side ovine BLG standards. Reduced transgenic animals from PLC A lines showed no detectable levels of protein by Coomassie staining (figure 5.21) or silver staining (data not shown).

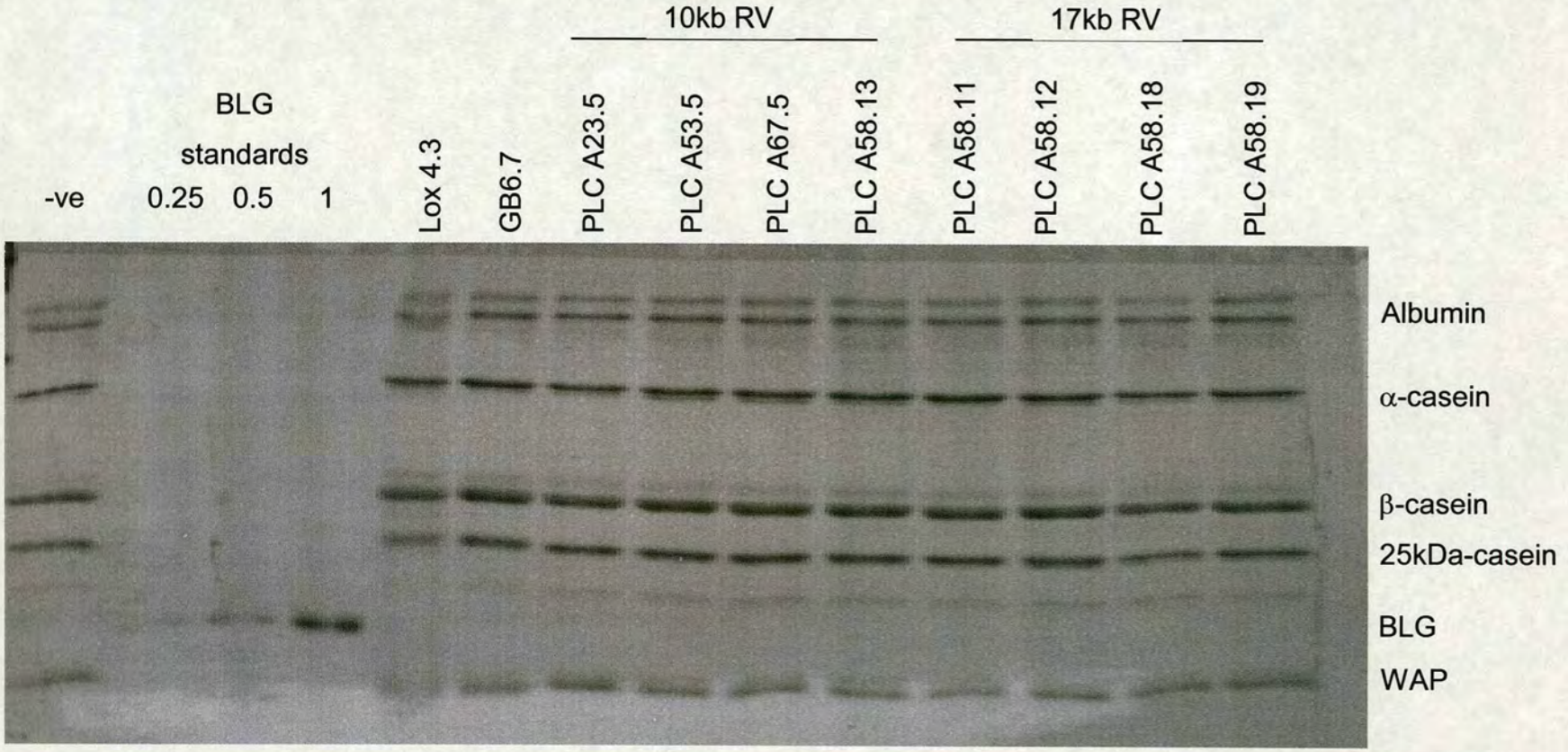


Figure 5.21. SDS-PAGE analysis of PLC A milk proteins. Mouse milk samples were electrophoresed in 18% polyacrylamide gels under reducing conditions and stained with Coomassie Blue. -ve indicates a 1/250 dilution of defatted control non-transgenic mouse milk sample. 1/250 dilutions of defatted milk from from either 10kb RV or 17kb RV locus transgenic animals was loaded along side standards of BLG.

As discussed in Chapter 4 the lowest level of protein detectable by Coomassie staining is approximately 0.1μg, while the lowest detection level of silver staining is approximately 1 –10 ng (Harlow and Lane 1988).

Western blots were carried out to determine if any detectable protein was present in the milk samples from the reduced animals. No protein was found in any 10kb or 17kb RV locus animal even when a 1/50 dilution of milk was used (figure 5.22). Therefore if any BLG protein was present it must be below the detection level of Western blotting (approximately 1ng; Harlow and Lane 1998).

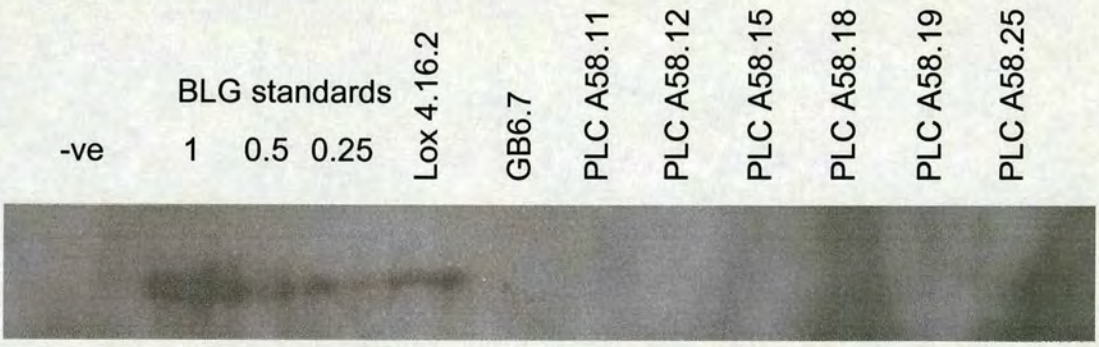


Figure 5.22. Western blot of milk proteins from 17kb RV locus animals. 1/50 dilutions of milk protein samples were run on a SDS-PAGE gel, electroblotted to nitrocellulose and probed with rabbit anti – BLG and then HRP conjugated anti rabbit IgG. The BLG standards are in μg.

5.3 CONCLUSIONS

The microinjection of Cre recombinase had led to a decrease in the number of copies of the transgene throughout the animal. In animals from the lox 4 lines there were obvious 'reduced arrays' corresponding to a 10kb RV locus, a 17kb RV locus and larger arrays. When animals were bred for either the 10kb RV locus or the 17kb RV locus no 'parental array' was found, indicating that these G1 animals carried only one type of array in all their cells. The PGK-Cre construct was not detected in any of the 'reduced array' animals indicating it had failed to integrate.

The multiple probing of EcoRV southern blots showed that both the 10kb and 17kb RV loci hybridised to the promoter and coding sequence of the transgene. This suggested that there was an intact copy of the BLG-loxP transgene in both 'reduced arrays'. When SspI digests were carried out, the 17kb RV locus showed that it had an intact repeat fragment. Therefore as far as could be shown by Southern blotting, the 17kb RV locus had an intact transgene fragment.

Although the 10kb RV locus hybridised to all the transgenic probes used, the Southern blotting data could not confirm the presence of an intact transgene in these animals. The remaining fragment after Cre recombination (6kb by SspI disgestion) hybridised to both the 5' probe and an internal probe. This fragment is large enough to include all the transgene from its 5' end to the SspI site (the distance between the 5' end and the SspI site is 5.4kb). The failure to produce a 7kb product from PCR of the 10kb RV locus

may indicate that there are rearrangements or deletions of the transgene, suggesting that the 10kb RV locus has been functionally inactivated. The 17kb RV locus does however, produce a 7kb PCR product which along with the Southern blotting data, strongly suggests that these animals contain at least one intact transgene.

In situ of these 'reduced array' animals showed that after the reduction in copy number there was a reduction in the percentage of cells that expressed in the mammary gland. Protein analysis failed to detect any protein from either the 10kb or 17kb RV locus animals indicating that the percentage of the mammary gland expressing protein is below available detection limits.

The 10kb RV locus animals only showed isolated dots of expression by *in situ* hybridisation, which were never found consistently between sequential sections. This was assumed to represent background as an artefact of the *in situ* technique rather than genuine expression. When taken in context with the failure to PCR a 7kb fragment, the most plausible explanation for the results in this chapter is that the 10kb RV locus does not contain a functional intact transgene. This also agrees with the data in Chapter 4, which failed to find any expression from the lox 4/Cre 10kb array animals.

In constrast, the 17kb RV locus animals showed discrete patches of expression which could be shown to exist across sequential sections. Therefore as these animals did express in some cells, they must contain a

functional transgene. The failure to find any protein expression or mRNA from the total mammary RNA indicates that although the 17kb RV locus can express, it is expressing at a level below the detection limit of these techniques. This result shows that when a multicopy BLG array is reduced to two copies the expression has also decreased. In Chapter 4, a model was proposed whereby the timing of deletion could be too late to rescue the transgene array. As this deletion occurred at a very early timepoint in development it is unlikely that the reduction in expression is due to the timing of deletion being too late.

The 17kb RV locus animals are dramatically different from their parental line, lox 4. The parental lox 4 animals express in approximately 50% of the alveoli examined, while the 17kb RV locus animals express in less than 10% of the alveoli examined. This is therefore counter to the hypothesis that the high copy number transgenes can be responsible for variegation. In this example, the high copy number lines express more efficiently than low copy number lines.

The fact that the 17kb RV locus does express the transgene but only in very discreet patches of the mammary suggests there maybe a threshold of copies required for uniform expression. This supports the model put forward in Chapter 4, that there is a 'buffer' effect allowing some transgenes within a multicopy array to escape silencing effects from the surrounding genome. If the number of copies is reduced, the buffer is removed leading to more cells failing to express the transgene. More PLC A lines are being bred currently that contain higher copy arrays (possibly 3 and 4 copies) which

would be of great interest to see if there is indeed a threshold above which uniform expression could be expected.

Alternatively, if there is inactive version of BLG contained within the 17kb RV locus it may act as a silencing element. If this inactive copy was targeted by silencing machinery, then it may have a direct effect on the full 7kb transgene associated with it. The proximity of the two may mean that in the majority of cells the functional 7kb has be repressed by the copy before it. It may be targeted to a repressive domain within the nucleus or maybe condensed into heterochromatin due to the interaction of the inactive copy and silencing machinery. The only way to determine if the inactive copy silences the 7kb would be to isolate the functional 7kb from the inactive version and determine if without the 'silencing' action if it can express in a higher proportion of the cells.

CHAPTER SIX

DISCUSSION

6.1 INTRODUCTION

The advent of pronuclear microinjection was a major development in molecular biology. Transgenic animals are now common tools to investigate many biological questions (Petters and Sommer 2000). It has become clear however, that not every transgene will express in the expected manner. Many factors influence the expression of the transgene: from those intrinsic to the construct i.e. enhancers, to those extrinsic to the construct which includes a wide variety of genome effects (as discussed in Chapter 1). In some lines of mice variegation reminiscent of PEV in *Drosophila* had been found as a consequence of the genomic location of the transgene (Dobie *et al* 1996). These position effects, whereby the surrounding genome can stochastically silence a gene were proposed to occur by a spread of heterochromatin from silenced regions to the gene (Henikoff 1990).

This project was initiated in mice to determine the relationship between the copy number of a transgenic array and its expression profile. Specifically, did high copy arrays lead to stochastic silencing of the transgene in different cells? The use of a site-specific recombinase, allowed a direct comparison of different copy number arrays at the same genomic location. If the hypothesis that the repeat nature of the high copy array induced a

genomic silencing effect, then upon reduction of copy number there would be expected to be a more uniform pattern of expression. From the results in this thesis, using a loxP containing BLG transgene, this was not the case. After reduction in copy number, the number of cells within the mammary gland that expressed was reduced.

6.2 BLG TRANSGENES AND COPY NUMBER

Many lines of mice have been created that carry a BLG transgene (Simons *et al* 1987, Harris *et al* 1991, Whitelaw *et al* 1992). In many cases it was established that there was a link between copy number and the levels of expression, and that the more copies present within an array the higher the BLG protein levels within milk (Whitelaw *et al* 1992). In the early experiments BLG expression levels were determined by the BLG protein/mRNA levels within milk or mammary gland RNA samples, and did not address the pattern of cellular expression.

When the cellular patterns of three BLG lines were investigated, two lines showed discreet patches of expression (Dobie *et al* 1996). This variegation within a tissue was thought to be due to the position within the genome of these two lines, i.e. centromeric, which was close to facultive heterochromatin. It was however noted that these two lines were also high copy number, and so this effect could not be ruled out. The two variegating lines 7 and 45 had approximately 25 and 17 copies each.

6.3 BLG-LOXP LINES

To determine what effect different transgene copy numbers had at the same integration site, BLG transgenic lines that could be manipulated by Cre recombination were designed. In chapter 3, five high copy lines containing a BLG-loxP transgene were generated. The expression profile was analysed using *in situ* hybridisation, Northern blots and milk composition analysis. Two of the lines showed patchy BLG mRNA expression in the mammary gland. The other three lines showed a BLG mRNA pattern consistent with an endogenous gene, β-casein. The lines with patchy expression showed variation in the levels of BLG protein found in the milk. This variation in milk protein correlates with the variation between animals, in the number of cells that express the transgene in the mammary gland.

Comparison between the five lines cannot be drawn as the integration site of each will be different, given the randomness of integration using pronuclear microinjection. However it is interesting that three of the five BLG-loxP lines chosen for their high copy arrays showed uniform expression. This indicates that BLG high copy number lines are in some cases able to express efficiently. The location of the transgene within the genome of these lines is obviously permissive to uniform expression, regardless of the copy number of the array. Other examples of genomic locations permissive of transgene expression, regardless of the structure or the transgene have been reported (Wallace *et al* 2000).

One issue that was not addressed in this study was the possibility that genetic modifiers of variegation exist within the genetic backgrounds of the mice. The original BLG-loxP construct was injected into hybrid mice. The BLG-loxP animals are therefore not all on the same genetic background. If one of the strains used to create the hybrids contained a genetic modifier locus this would segregate to different animals. From the small number of mice used, no segregation between uniform and variegating animals was found within a lox line. However, modifiers may be expected to affect the levels of variegation and no extensive backcrossing was done to determine if the loxP lines were influenced by modifiers. A previous study on BLG variegating lines found no evidence that the genetic background had an effect on the variegation (Dobie *et al* 1996).

One limitation within the project was the use of quantification software to determine the copy number of the lines and the protein expression levels. As the project relied on the generation of high copy number lines to start with, the estimation of copy number was important. Several areas may lead to problems: the use of external controls may lead to pipetting errors, the accuracy of quantification software is determined by the ability to accurately box the data points, the sensitivity of the phosphor screens, and whether the samples fall within the dynamic range of the screen. Alternative methods exist to determine copy numbers of transgenes, i.e. pulse field gel electrophoresis. The use of PFGE for copy number is based on size of the fragments, not the intensities so it avoids these problems. Protein expression levels were determined from multiple sampling to try and reduce any human errors in boxing data points.

6.4 TRANSGENE REDUCTION IN THE MAMMARY GLAND

In chapter 4, BLG-loxP lines 4, 5 or 9 were crossed with a BLG-Cre line to produce a mammary-specific reduction in copy number. The mammary gland contains more than one cell type, and the BLG promoter used to drive Cre recombinase is expressed only in the secretory epithelial cells (Harris *et al* 1991). A major drawback to this method was that the mammary gland was then mosaic for two types of transgenic array; parental and reduced. From Southern blotting of the double transgenic animals all the lox lines showed a reduction in copy number based on the intensity of the 'parental array'. However, only lox 4/Cre showed a discrete restriction fragment smaller than the 'parental array'. When analysed further using multiple probes, this reduced fragment appeared to comprise an intact transgene. However, the Southern blotting data could not exclude the possibility that the transgene had been rearranged, making it functionally inactive.

No reduced fragment could be found in the double transgenic animals of the other two lines, even though the intensity of the parental array was significantly reduced. It is possible that reduced but not single copy arrays were trapped within the parental arrays due to the percentage agarose gels used. The remaining parental arrays would be derived from non Cre-expressing cells. This would suggest that the Cre expression was not fully penetrant within the mammary gland. The BLG-Cre line had been reported to express in a uniform manner in the secretory epithelial cells, although the original studies suggested it only expressed in 80% of the gland (Selbert *et al* 1998). It did not formally exclude the possibility that a small fraction of the secretory epithelial cells failed to undergo Cre mediated recombination.

The inability to distinguish between parental and reduced arrays due to their shared sequence precluded further molecular analysis.

When double transgenic animals were analysed by *in situ* hybridisation, it was clear that in all three lines crossed to a Cre recombinase expressing line there had been a substantial reduction in the number of cells within the mammary gland that expressed the transgene. There was also a correlation between the reduction in copy number and a reduced level of BLG protein found within the milk.

If the Southern blotting and *in situ* results are examined together, the data from the lox 5 and 9 double transgenic animals suggests that the areas of the mammary gland that still express the transgene, are due to non Cre-expressing cells still containing the parental arrays. In the lox 4 double transgenics, on the basis of Southern blotting data, there appeared to be an intact transgene after reduction. However, given the fact that no lox 4/Cre animal showed expression means that the 'reduced array' failed to express or that it had been inactivated in some way.

If the transgenes had not been inactivated or deleted by the reduction, then two other possible explanations could explain the decrease in expression after reduction in copy number. One model was that the deletion had occurred too late to rescue the expression. This was based on the fact that the mammary gland is an unusual tissue; it is only fully differentiated in adult females after pregnancy. The BLG-Cre transgene is expressed co-ordinately with other milk protein genes and only reaches its maximum

expression during lactation (Harris *et al* 1991, Selbert *et al* 1998). The Cre mediated recombination in these animals would therefore only occur in the differentiated mammary tissue. The reduction in copy number could therefore occur after a developmental mark had silenced the transgene, and after this point the modification of copy number would have no effect. Interestingly, the patterns of expression in these variegated lines and others previously studied (Dobie *et al* 1996) suggests that the variegation may be clonally derived. By the intensity of *in situ* hybridisation, the expression did not appear to vary in the level per cell, only the number of cells expressing varied. This may suggest that during differentiation, a process of commitment to either an active or silent state is undertaken and that this is then maintained during the further growth and development of the mammary gland. Such patterns would indicate that the decision to silence an array may be taken at an early developmental stage. An alternative model was that a 'buffer' zone in the multicopy arrays allowed some transgenes within that array to escape genomic silencing effects. After reduction this 'buffer' zone would had been removed, making the remaining transgenes more prone to silencing.

6.5 TRANSGENE REDUCTION BY MICROINJECTION

Having created a reduction of copy number in the mammary gland, the same three lox lines were then used to create a reduction by microinjection of a PGK-Cre construct. This reduction was therefore much earlier in development than the mammary gland reduction discussed in Chapter 4. This method provided animals that were not mosaic for two transgenic arrays within the same tissues, and allowed lines to be established.

Clearer analysis could then be carried out without the interference from non-reduced cells, unlike the situation in Chapter 4.

Southern blotting again failed to show any evidence for a identifiable 'reduced array' in the lox 9 animals. All the animals analysed showed a smear of reduced arrays. From the lox 5 line only one animal was detected carrying a 'reduced array'. The rest showed smears indicating that the transgene had undergone Cre mediated recombination to varying degrees in different cells. Taken together with the results from Chapter 4 it seems likely that these two transgenic loci are unsuitable for recombination to a single intact transgene. Alternatively, it maybe that the Cre recombinase is less efficient at recombining the structures of these loci, given the complex patterns of fragments found in the lox 9 animals after microinjection of Cre.

By constrast, microinjected lox 4 animals did produce identifiable 'reduced arrays' by EcoRV digestion. The smallest of these arrays co-migrated with the 10kb 'reduced array' from the double transgenics of Chapter 4. The next sized array was estimated to be 17kb by EcoRV digestion, which suggests an increase in copy number of one given that the transgene is 7kb long. Other arrays were larger and were proposed to carry more transgenes. Interestingly, the two reports in the literature that specifically used Cre recombinase to reduce transgene copy number both resulted in two distinct arrays after recombination from the one parental line (Garrick *et al* 1998, Ramirez *et al* 2001). Two arrays are consistent with only partial deletion in some cells. Thus Cre recombinase may frequently catalyse different amounts of recombination in different cells of the same animal.

Further Southern blotting was used to examine the structure of the 'reduced array'. It was established that the 10kb RV locus did not contain a repeat length. The 10kb RV locus from both lox4/BLG-Cre and the microinjected Cre failed to produce a 7kb PCR product. Although the Southern blotting data suggested the 10kb RV locus contained both promoter and coding sequence, the PCR result suggests it may not be in the correct arrangement, although further work is required to confirm this. The failure to find expression from this array in either experiment, therefore, is most likely due to inactivation of the transgene during recombination.

The 17kb RV locus contained a repeat length by SspI digestion indicating that the 'reduced array' had a minimum of two transgene copies in a head to tail repeat, and confirmed that it had at least one intact transgene. The 17kb RV locus also produced a 7kb PCR product. This correlates with the SspI digestion which showed a repeat fragment of similar intensity to the junction fragment, suggesting that there is only one repeat fragment present within the 17kb RV locus.

In situ data from the 10kb RV locus showed very infrequent dots of hybridisation. However, these were extremely rare, and were not shown to be consistent between sequential sections. The *in situ* data showed no genuine expression, which when considered in parallel to the Southern blotting and PCR data support the theory that the 10kb RV locus did not express because it was in some way functionally inactivated after recombination.

In contrast, the 17kb RV locus showed distinct patches of expression which were found in all the animals examined and were consistent between sequential sections. The level of expression however, was dramatically different to the parental lox 4 array. The multicopy lox 4 animals expressed in approximately 50% of the alveoli, while the 17kb RV locus animals expressed in less than 10% of the alveoli. Although, Western blotting is able to detect approximately 1-10ng of protein (Harlow and Lane 1988), no BLG protein was found in the 17kb RV locus animals. This compares with the original lox 4 animals that had a mean BLG protein level of 1.2 mg/ml. The level of expression had therefore, decreased to a point undetectable by Western blotting or Northern blotting.

Real time PCR may be one technique that could determine gene expression from the 17kb RV locus animals, given the failure to find any RNA by Northern blotting. Real time PCR allows the quantification of gene expression by the use of fluorescent probes that are in direct proportion to the amount of PCR product in a reaction. In a PCR reaction the amplification initially proceeds at an exponential rate, moving into a linear phase followed by a plateau phase. Real time PCR quantitates the amount of product produced every cycle, allowing quantification over a broad range. The use of different fluorescent reporters for different genes allows multiplex PCR to be carried out in a single tube. Real time PCR can therefore be used to compare expression from a transgene relative to an endogenous gene within the same biological sample. Given the sensitivity of PCR and the ability to quantitate the products, it could be used as an alternative method to *in situ* hybridisation for quantifying BLG transgene expression.

It is clear from the 17kb RV locus animals, that a reduction in copy number led to a decrease in expression, due to a decrease in the number of cells that expressed the transgene. From these results, at this genomic location a high copy number (16-18 copies) of BLG transgenes actually express significantly more efficiently than a lower copy number (one copy). This contradicts the original hypothesis of the project and shows that the variegation of this BLG line was not directly related to its high copy number.

6.6 OTHER TRANSGENE REDUCTIONS

Previous investigations into copy number effects have provided conflicting reports. In one case, reduction of copy number led to an increase in the number of cells that expressed the transgene (Garrick *et al* 1998). This study used a hybrid construct which contained an α-globin promoter driving a *lacZ* reporter and the α-HS 40 enhancer-like element. This construct previously showed variegated expression and was not thought to express in a copy number dependent manner (Robertson *et al* 1995). When this transgene was reduced by microinjection of Cre recombinase, in one line there was a 1000-fold increase in the number of cells expressing. Another line showed a 180-fold increase after reduction to one copy. Therefore, at two genomic locations with the α-globin transgene, high copy arrays were directly associated with silencing of expression. A number of differences exist with this case and the examples in this thesis. The globin transgene contains part of the globin locus and an enhancer. Given the complexity of the natural globin locus these sequences when present as single copies may be unusually dominant in controlling the effects of the surrounding genome. As discussed in Chapter 1 variegation can be thought of as a complex result

of the effects from the position within the genome, the repeated nature of the transgene and the intrinsic elements of the transgene. It is possible that after removal of effects due to the repeated nature of the transgene, that this α-globin transgene contains elements strong enough to overcome only position effects, allowing it to express in a position independent manner.

In the other case, reduction in copy number had no effect on the expression profile of the transgene (Ramirez *et al* 2001). In this study a Keratin *K5* construct was linked to the *lacZ* reporter. The construct was very prone to variegation, indeed in the 6 lines used to study copy number, two were silenced and the other four variegated. This is unlike BLG transgenes, which of the eight lines studied at a cellular level only half displayed variegated expression, while the other half were uniformly expressed. If the *lacZ* reporter was replaced with other sequences (i.e. Human epidermal growth factor reporter; HEGFR) uniform expression occurred using the same Keratin *K5* regulatory sequences. In this study, no quantification of variegation before and after Cre reduction was carried out. This study also did not state if the transgene after Cre recombination was examined to prove it was unrearranged by the reduction. The copy number reduction showed no effect on the variegation of the transgenes. Given the known silencing effects of *lacZ* on adjacent sequences (Cohen-Tannoudji *et al* 2000; Clark *et al* 1997 also show silencing from other reporter sequences), it seems likely that the continued variegation after reduction, was actually due to the presence of the *lacZ* sequence. It maybe that the *lacZ* has such a strong silencing effect on the keratin construct that even when any negative effects from repeat structures are removed the transgene is still variegated. Indeed, when a single *lacZ* insertion was targeted to the mouse actin locus, this resulted in

highly variegated expression of the targeted allele (Prof A.J. Clark; pers comm.)

An important distinction between both previous published studies and this project is that the BLG transgene investigated is not a hybrid construct and it did not contain a reporter sequence. This allows the BLG transgene to be examined without the influence of inherent silencing effects from reporter sequence (i.e. from *lacZ*) within the construct. Again in contrast, BLG does not contain an LCR or specific enhancers, which may make it more prone as a single copy to position effects than the α-globin single copy.

From the small number studies into repeat induced silencing in transgenic mice, it is clear that for some transgenic constructs repeat structures should be avoided i.e. the globin construct (Garrick *et al* 1998), while for others the construct may only efficiently express when present as a multicopy array, i.e. the BLG construct.

6.7 BUFFER MODEL OF VARIEGATION

From the results in this thesis, silencing is highest among low copy number BLG arrays. In Chapter 4, two models were proposed to account for the reduction in expression after reduction in copy number. The reduction in the mammary gland was proposed to be too late in development to rescue the transgene from a silencing mark. However, the results from chapter 5

show that even if a reduction in copy number occurs very early in development, a reduction in expression is still seen. This makes the model proposing a developmental mark seem unlikely. The other model described in Chapter 4 whereby a buffer zone in multicopy arrays exists can explain the reduced expression. Having large numbers of a repeat allows some transgenes in the array to escape silencing, due to them being buffered against genome effects due to the large amount of copies that the genome must inactivate (figure 6.1). After reduction there would be no protection from the surrounding genome for the reduced array, therefore reducing the number of cells that manage to escape this genomic effect. Given that PEV in *Drosophila* is accompanied by heterochromatinization of transgenes (Dorer and Henikoff 1994) the large arrays of BLG may provide a buffer zone against this.

The unmodified BLG gene has previously been used to rescue expression from poorly expressed hybrid constructs by co-microinjection (Clark *et al* 1992). This resulted in a high proportion of the founders that comprised both constructs at the same integration site. This led to a significant increase in the frequency of expression from the hybrid constructs. It was thought that the increase was due to changes associated with activation or transcription of the unmodified BLG spreading to the hybrid transgene. However, later work showed that the rescue effect was associated with transcription of the BLG, but not its translation (Yull *et al* 1997). It was thought this represented the changes in chromatin structure round the BLG gene when it activates transcription spreading to the hybrid gene (Whitelaw 1995). Therefore it may be that the buffer zone described

here is also due to the changes in chromatin structure surrounding an active copy, being able to withstand the influences of the genome effects.

An alternative theory is that the 17kb RV locus contains a inactive version of BLG that has an inhibitory effect on the remaining intact copy associated with it. Just as BLG can rescue poorly expressed transgenes, it may be possible that the inactive version is acting as a negative element attracting silencing machinery, which in turn reduces/abolishes the expression from the intact copy. To formally prove that the inactive copy is acting as a silencer, the 17kb RV locus would need to be cloned out from the mouse DNA, the two transgene copies separated and the intact copy used on its own at single copy to determine if the inactive version was indeed responsible.

The availability of higher copy number BLG arrays (possibly 3-4 copies) would allow investigations into whether there is a threshold for efficient expression at these locations. It must be considered that in other genomic locations low copy number BLG arrays can express (<5 copies gave uniform expression; Dobie *et al* 1996). Therefore it may be that for the genomic location of the array studied in this project, the position effects from the surrounding genome are strong and that only high copy arrays can escape from this by 'buffering' the internal copies.

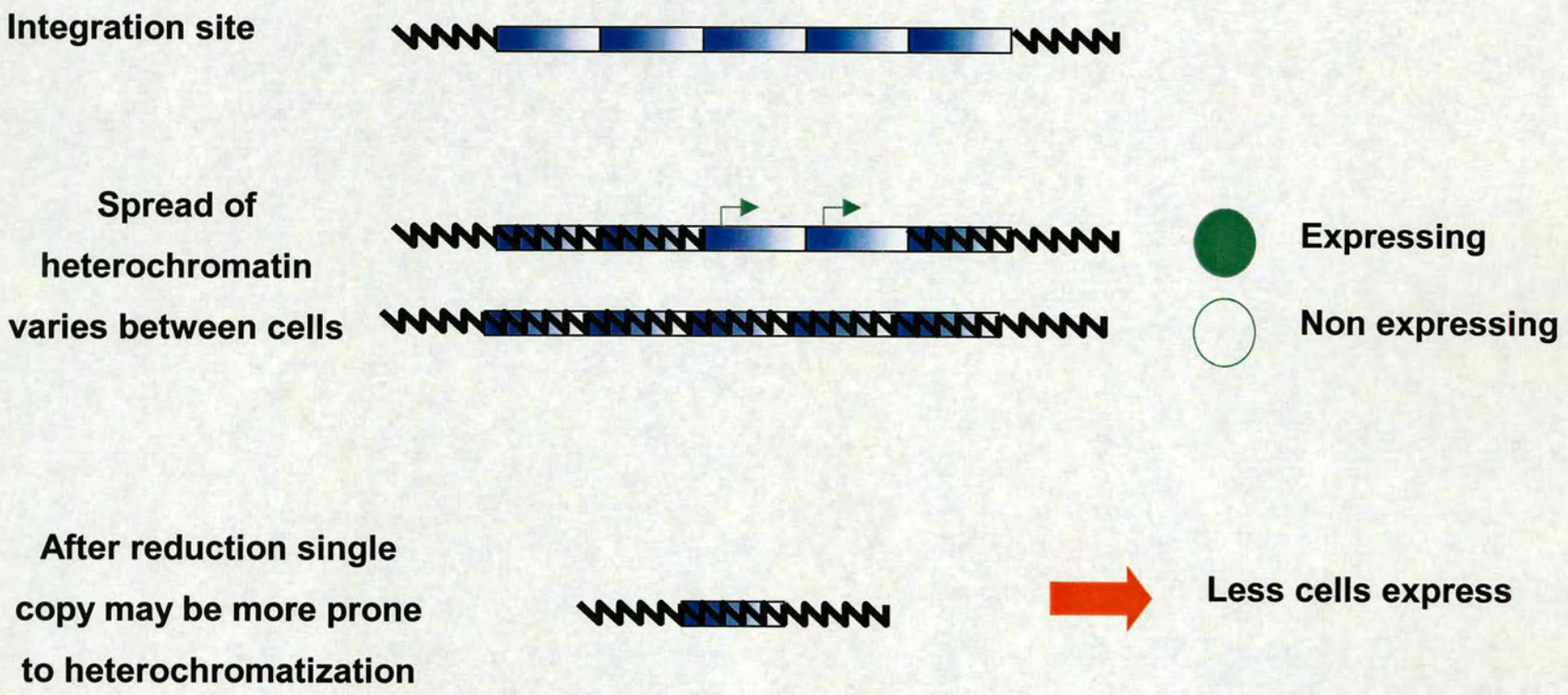


Figure 6.1. Model of buffer effects from multiple copies. Having a high copy array would allow some transgenes within the array to escape the effects of the surrounding genome, leading to some expression in the tissue in a patchy manner. After reduction, this 'buffer' would be lost and so fewer cells would escape the genome effects.

Cre recombinase has been extensively used for genetic engineering in transgenic mice and was thought to catalyse base pair precise recombinations (Kilbey *et al* 1993). However, in this study it use appears to have caused in some cases partial deletion within an array or specific deletions. Previous use of Cre recombinase to reduce copy number also produced rearrangement of a reduced single copy transgene (Garrick *et al* 1998). It can be difficult to accurately determine the exact nature of transgenic loci and they can often integrate with duplications or filler sequence (Bishop 1997). It is clear that care must be taken to determine the structure of any array to be used with Cre recombinase, if they do not contain precise head to tail repeats unwanted recombination events may occur.

It has been proposed that Cre recombinase be routinely used to reduce transgene copy number in attempts to avoid transgene silencing (Jasin *et al* 1996, Metzger and Feil 1999). Given the results from this thesis, reducing transgene copy number may not lead to improved transgene expression, and as with the example of BLG may indeed make the situation worse. The widespread use of Cre recombinase to reduce transgene copy number could also result in large numbers of rearranged transgenes, which may be functionally inactive. It is clear that great care must be taken in determining the structure of any 'reduced arrays' before conclusions are drawn. If the array does not contain exact head to tail arrays then the Cre catalysed recombination may give unwanted results. It is therefore possible that the use of Cre recombinase to reduce copy number on a routine basis may not be of benefit to improving transgene expression.

Variegation of transgene expression can be due to position effects or the structure of the array but it is probably best explained as a combination of the two, in that the structure of the array fails to overcome the position effects that it encounters. Therefore both the position and copy number are important factors in determining the expression of a transgene. From this study, it is clear that reduction of a high copy number line does not necessarily result in an increase in expression from a site of integration in the genome, which is subject to silencing. In fact, reduction can produce significantly more silencing. Repeat Induced Gene Silencing may therefore be due to the combination of location, copy number and the transgene itself, and does not apply to all high copy number lines. RIGS cannot therefore, be used to explain all the cases of variegated transgene expression in high copy number lines. In this study there appears to be a buffer effect, which has not been found with other transgenes, nevertheless, this effect may not be limited to BLG transgenes. In some cases, the repeated structure may allow a number of transgenes within an array to avoid silencing effects.

APPENDIX ONE

Conversion of rpm into RCF

Relative Centrifugal force (RCF; measured in x *g*) refers to the force during centrifugation that moves the particulate outwards from the centre of rotation. The force is proportional to the speed (rpm) squared and the rotor radius. It is calculated by the following equation:

$$RCF = 1.12\ r\ (rpm/1000)^2$$

Where:

r = the rotating radius between the particle being centrifuged and the axis of rotation in mm. r_{max} values are equal to the distance between the axis of rotation and the bottom of the centrifuge tube as it sits in the rotor. The r_{max} can be obtained from the manufacturer.

rpm = revolutions per minute (speed).

Some examples of frequently used speeds on centrifuge rotors are shown in table A.1.

Rotor	*r in mm*	*rpm*	*RCF*
Desktop microcentrifuge	83	13000	15700
Sorval SS 34 rotor	107	8000	7670
Sorval F16 rotor	88	5000	2464
		7000	4829
Jouan CR3000 Swing bucket rotor	127	4500	2880
Eppendorf 5810 R Swing bucket rotor	185	4000	3220

Table A.1. Conversion table for rpm into RCF.

REFERENCES

Adra, C.N., Boer, P.H. & McBurney, M.W. (1987). Cloning and expression of the mouse *pgk-1* gene and the nucleotide sequence of its promoter. *Gene* **60**, 65-74.

Al-Shawi, R., Kinnaird, J., Burke, J. & Bishop, J.O. (1990). Expression of a foreign gene in a line of mice is modulated by chromosomal position effect. *Mol Cell Biol* **10**, 1192-1198.

Ali, S. & Clark, A.J. (1988). Characterisation of the gene encoding ovine β-lactoglobulin. *Journal of Molecular Biology* **199**, 415-426.

Allen, D.N., Norris, M.L. & Surani, M.A. (1990). Epigenetic control of transgene expression and imprinting by genotype-specific modifiers. *Cell* **61**, 853-861.

Anandalakshmi, R., Pruss, G.J., Ge, X., Marathe, R., Mallory, A.C., Smith, T.H. & Vance, V.B. (1998). A viral suppressor of gene silencing in plants. *Proc Natl Acad Sci USA* **95**, 13079-13084.

Angerer, L.M. & Angerer, R.C. (1992). In situ hybridisation to cellular RNA with radiolabelled RNA probes. In *In situ Hybridisation: A Practical Approach* (Wilkinson, D.G., ed.), Oxford University Press, Oxford, pp. 15-31.

Araki, K., Araki, M., Miyazaki, J.-I. & Vassalli, P. (1995). Site specific recombination of a transgene in fertilized eggs by transient expression of cre recombinase. *Proc Natl Acad Sci USA* **92**, 160-164.

Archibald, A.L., McClenaghan, M., Hornsey, V., Simons, J.P. & Clark, A.J. (1990). High-level expression of biologically active human α1-antitrypsin in the milk of transgenic mice. *Proc Natl Acad Sci USA* **87**, 5178-5182.

Assaad, F.F., Tucker, K.L. & Signer, E.R. (1993). Epigenetic repeat-induced gene silencing (RIGS) in *Arabidopsis*. *Plant.Mol.Bio.* **22**, 1067-1085.

Baubonis, W. & Sauer, B. (1993). Genomic targeting with purified Cre recombinase. *Nucleic Acids Res.* **9**, 2025-2029.

Baulcombe, D.C. (1996). RNA as target and initiator of post-transcriptional gene silencing in transgenic plants. *Plant.Mol.Bio.* **32**, 79-88.

Bell, A.C. & Felsenfeld, G. (2000). Methylation of a CTCF-dependent boundary controls imprinted expression of the *Igf2* gene. *Nature* **405**, 482-485.

Belyaeva, E.S., Demakova, O.V., Umbetova, G.H. & Zhimulev, I.F. (1993). Cytogenetic and molecular aspects of position-effect variegation in *Drosophila melanogaster*. V. Heterochromatin-associated protein HP1 appears in euchromatic chromosomal regions that are inactivated as a result of position-effect variegation. *Chromosoma* **102**, 583-590.

Berg, P. (1981). Dissections and reconstructions of Genes and Chromosomes. *Science* **213**, 296-303.

Bernardi, G. (1995). The human genome: Organisation and evolutionary history. *Annual Rev.Genet.* **29**, 445-476.

Bestor, T.H. (1998). Methylation meets acetylation. *Nature* **393**, 311-312.

Bhattacharya, S.K., Ramchandani, S., Cervoni, N. & Szyf, M. (1999). A mammalian protein with specific demethylase activity for mCpG DNA. *Nature* **397**, 579-583.

BioRad. (2000). Quantity One User Guide for Version 4.

Birchler, J.A., Pal-Bhadra, M. & Bhadra, U. (2000). Making noise about silencing: repression of repeated genes in animals. *Current Opinion in Genetics and Development* **10**, 211-216.

Birchler, J.A., Bhadra, U., Pal Bhadra, M. & Auger, D.L. (2001). Dosage-dependent gene regulation in multicellular eukaryotes: implications for dosage compensation, aneuploid syndromes, and quantatitive traits. *Dev.Biol.* **234**, 275-288.

Bird, A. & Wolffe, A.P. (1999). Methylation-Induced Repression - Belts, Braces, and Chromatin. *Cell* **99**, 451-454.

Bird, AP. (1995). Gene number,noise-reduction and biological complexity. *Trends Genet.* **11**, 94-100.

Bishop, J.O. & Smith, P. (1989). Mechanism of chromosomal integration of microinjected DNA. *Mol.Biol.Med.* **6**, 283-293.

Bishop, J.O. (1997). Chromosomal insertion of foreign DNA. In *Transgenic Animals Generation and Use* (Houdebine, L.-M., ed.), Harwood Academic Publishers, Amsterdam, pp. 219-223.

Bonifer, C., Jägle, U. & Huber, M.C. (1997). The chicken lysozyme locus as a paradigm for the complex developmental regulation of eukaryotic gene loci. *J.Biol.Chem* **272**, 26075-26078.

Boyd, A.L. & Samid, D. (1993). Molecular biology of transgenic animals - Review. *J.Animal Sci.* **71** , 1-9.

Boyes, J. & Bird, A. (1991). DNA methylation Inhibits Transcription Indirectly via a Methyl-CpG Binding Protein. *Cell* **64**, 1123-1134.

Bradl, M., Larue, L. & Mintz, B. (1991). Clonal coat color variation due to a transforming gene expressed in melanocytes of transgenic mice. *Proc Natl Acad Sci USA* **88**, 6447-6451.

Bradley, A., Evans, M., Kaufman, M.H. & Robertson, E. (1984). Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. *Nature* **309**, 255-256.

Brinster, R.L., Chen, H.Y. & Trumbauer, M. (1981). Somatic expression of Herpes Thymidine Kinase in Mice following Injection of a Fusion Gene into Eggs. *Cell* **27**, 223-231.

Brinster, R.L., Chen, H.Y., Traumbauer, M.E., Yagle, M.K. & Palmiter, R.D. (1985). Factors affecting the efficiency of introducing foreign DNA into mice by microinjecting eggs. *Proc Natl Acad Sci USA* **82**, 4438-4442.

Bronson, S.K., Plaetn, E.G., Kluckman, J.R., Hagaman, J.R., Maeda, N. & Smithies, O. (1996). Single copy transgenic mice with chosen site integration. *Proc Natl Acad Sci USA* **93**, 9067-9072.

Brooks, A.R., Nagy, B.P., Taylor, S., Simonet, W.S., Taylor, J.M. & Levy-Wilson, B. (1994). Sequences containing the second-intron enhancer are essential for transcription of the human apolipoprotein B gene in the livers of transgenic mice. *Mol Cell Biol* **14** , 2243-2256.

Brown, C.J., Ballabio, A., Rupert, J.L., Lafreniere, R.G., Grompe, M., Tonlorenzi, R. & Willard, H.F. (1991). A gene from the region of the human x-inactivation center is expressed exclusively from the inactive x-chromosome. *Nature* **349**, 38-44.

Brown, C.J. & Willard, H.F. (1994). The human X-inactivation centre is not required for maintance of X-chromosome inactivation. *Nature* **368**, 154-156.

Brown, K.E., Guest, S.S., Smale, S.T., Hahm, K., Merkenschlager, M. & Fisher, A.G. (1997). Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. *Cell* **91**, 845-854.

Brownell, J.E. & Allis, C.D. (1996). Special HATs for special occasions: linking histone acetylation to chromatin assembly and gene activation. *Current Opinion in Genetics and Development* **6**, 176-184.

Bullock, D.W., Damak, S., Jay, N.P., Su, H.-Y. & Barrell, G.K. (1997). Improved wool production from Insulin-like Growth Factor 1 targeted to the wool follicle in transgenic sheep. In *Transgenic Animals: Generation and Use* (Houdebine, L.M., ed.), Harwood Academic Publishers, Amsterdam, pp. 507-509.

Bunting, M., Bernstein, K.E., Greer, J.M., Capecchi, M.R. & Thomas, K.R. (1999). Targeting genes for self-excision in the germ line. *Genes Dev.* **13**, 1524-1528.

Burdon, T.G., Maitland, K.A., Clark, A.J. & Watson, C.J. (1994). Regulation of the sheep β-lactoglobulin gene by lactogenic hormones is mediated by a transcription factor that binds an interferon-γ activation site-related element. *Mol Endocrinol* **8**, 1528-1536.

Butner, K. & Lo, C.W. (1986). Modulation of *tk* Expresssion in Mouse Pericentromeric Heterochromatin. *Mol Cell Biol* **6 (12)**, 4440-4449.

Camper, S.A. (1987). Research applications of transgenic mice. *Biotechniques* **5,** 638-650.

Capecchi, M.R. (1989). Altering the genome by homologous recombination. *Science* **244**, 1288-1292.

Caterina, J.J., Ciavatta, D.J., Donze, D., Behringer, R.R. & Townes, T.M. (1994). Multiple elements in human β-globin locus control region 5' HS 2 are involved in enhancer activity and position-independent, transgene expression. *Nucleic Acids Res* **22 (6)**, 1006-1011.

Chada, K., Magram, J., Raphael, K., Radice, G., Lacy, E. & Costantini, F. (1985). Specific expression of a foreign β-globin gene in erythroid cells of transgenic mice. *Nature* **314**, 377-380.

Choi, T., Huang, M., Gorman, C. & Jaenisch, R. (1991). A generic intron increases gene expression in transgenic mice. *Mol Cell Biol* **11**, 3070-3074.

Clark, A.J., Cowper, A., Wallace, R. & Simons, J.P. (1992). Rescuing transgene expression by co-integration. *Biotechnology* **10**, 1450-1454.

Clark, A.J., Archibald, A.L., McClenaghan, M., Simons, J.P., Wallace, R. & Whitelaw, C.B.A. (1993). Enhancing the efficiency of transgene expression. *Phil.Trans.R.Soc.Lond.B* **339** , 225-232.

Clark, A.J., Bissinger, P., Bullock, D.W., Dmak, S., Wallace, R.M., Whitelaw, C.B.A. & Yull, F.E. (1994). Chromosomal position effects and the modulation of transgene expression. *Reprod.Fert.Dev.* **6**, 589-598.

Clark, A.J., Harold, G. & Yull, F.E. (1997). Mammalian cDNA and prokaryotic reporter sequences silence adjacent transgenes in transgenic mice. *Nucleic Acids Res.* **25**, 1009-1014.

Clark, A.J. (1998). The mammary gland as a bioreactor: expression frequency and production of recombinant proteins. *Journal of Mammary Gland Biology and Neoplasia* **3**, 337-350.

Clark, A.J., Burl, S., Denning, C. & Dickinson, P. (2000). Gene targeting in livestock: a preview. *Transgenic Res.* **9**, 263-275.

Cockerill, P.N. & Garrard, W.T. (1986). Chromosomal loop anchorage of the kappa immunoglobulin gene occurs next to the enhancer in a region containing topoisomerase II sites. *Cell* **44**, 273-282.

Cogoni, C. & Macino, G. (1999). Gene silencing in *Neurospora crassa* requires a protein homologous to RNA-dependent RNA polymerase. *Nature* **399**, 166-169.

Cohen-Tannoudji, M., Vandormael-Pournin, S., Drezen, J.-M., Mercier, C., Babinet, C. & Morello, D. (2000). *lacZ* sequences prevent regulated expression of housekeeping genes. *Mechanisms of Development* **90**, 29-39.

Cranston, A., Dong, C., Howcroft, J. & Clark, A.J. (2001). Chromosomal sequences flanking an efficiently expressed transgene dramatically enhance its expression. *Gene* **269**, 217-225.

Davis, B.P. & MacDonald, R.J. (1988). Limited transcription of rat elastase I transgene repeats in transgenic mice. *Genes and Development* **2**, 13-22.

Dellaire, G. & Chartrand, P. (1998). Direct evidence that transgene integration is random in murine cells, implying that naturally occurring double-strand breaks may be distributed similarly within the genome. *Radiation Research* **149**, 325-329.

Denning, C., Burl, S., Ainslie, A., Braken, J., Dinnyes, A., Fletcher, J., King, T., Ritchie, M., Ritchie, W.A., Rollo, M., de Sousa, P., Travers, A., Wilmut, I. & Clark, A.J. (2001). Deletion of the α(1,3)galactosyl transferase (GGTA1) gene and the prion protein (PrP) gene in sheep. *Nature Biotechnology* **19**, 559-562.

deWit, T., Drabek, D. & Grosveld, F. (1998). Microinjection of Cre recombinase RNA induces site-specific recombination of a transgene in mouse oocytes. *Nucleic Acids Res.* **26**, 676-678.

Dobie, K.W., Lee, M., Fantes, J.A., Graham, E., Clark, A.J., Springbett, A.J., Lathe, R. & McClenaghan, M. (1996). Variegated transgene expression in mouse mammary gland is determined by the transgene integration locus. *Proc Natl Acad Sci USA* **93**, 6659-6664.

Dobie, K.W., Mehtali, M., McClenaghan, M. & Lathe, R. (1997). Variegated gene expression in mice. *Trends Genet.* **13**, 127-130.

Dorer, D.R. & Henikoff, S. (1994). Expansions of transgene repeats cause heterochromatin formation and gene silencing in Drosophila. *Cell* **77**, 993-1002.

Dorer, D.R. & Henikoff, S. (1997). Transgene repeat arrays interact with distant heterochromatin and cause silencing in *cis* and *trans*. *Genetics* **147**, 1181-1190.

Dorer, D.R. (1997). Do transgene arrays form heterochomatin in vertebrates? *Transgenic Res.* **6**, 3-10.

Eissenberg, J.C., James, T.C., Foster-Hartnett, D.M., Hartnett, T., Ngan, V. & Elgin, S.C.R. (1990). Mutation in a heterochromatin-specific chromosomal protein is associated with suppression of position-effect variegation in *Drosophila melanogaster*. *Proc Natl Acad Sci USA* **87**, 9923-9927.

Eissenberg, J.C. & Elgin, S.C.R. (2000). The HP1 protein family: getting a grip on chromatin. *Current Opinion in Genetics and Development* **10**, 204-210.

Ellis, J., Tan-Un, K.C., Harper, A., Michalovich, D., Yannoutsos, N., Philipsen, S. & Grosveld, F. (1996). A dominant chromatin-opening activity in 5' hypersensitive site 3 of the human β-globin locus control region. *The EMBO Journal* **15**, 562-568.

Evans, M.J. & Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154-156.

Eyestone, W.H. (1994). Challenges and progress in the production of transgenic cattle. *Reprod.Fert.Dev.* **6**, 647-652.

Falls, J.G., Pulford, D.J., Wylie, A.A. & Jirtle, R.L. (1999). Genomic Imprinting: Implications for Human Disease. *American Journal of Pathology* **154**, 635-647.

Feinberg, A.P. & Vogelstein, B. (1984). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Addendum.Anal.Biochem.* **137**, 266-267.

Festenstein, R., Tolaini, M., Corbella, P., Mamalaki, C., Parrington, J., Fox, M., Milou, A., Jones, M. & Kioussis, D. (1996). Locus control function and heterochromatin-induced position effect variegation. *Science* **271**, 1123-1125.

Festenstein, R., Sharghi-Namini, S., Fox, M., Roderick, K., Tolaini, M., Norton, T., Saveliev, A., Kioussis, D. & Singh, P. (1999). Heterochromatin protein 1 modifies mammalian PEV in a dose- and chromosomal-context-dependent manner. *Nat Genet* **23**, 457-461.

Flavell, R.B. (1994). Inactivation of gene expression in plants as a consequence of specific sequence duplication. *Proc Natl Acad Sci USA* **91**, 3490-3496.

Flower, D.R. (1996). The lipocalin protein family: Structure and function. *Biochem J.* **318**, 1-14.

Francastel, C., Walters, M.C., Groudine, M. & Martin, D.I.K. (1999). A Functional Enhancer Suppresses SIlencing of a Transgene and Prevents Its Localization Close to Centromeric Heterochromatin. *Cell* **99**, 259-269.

Franklin, G.C., Donovan, M., Adam, G.I.R., Holmgren, L., Pfeifer-Ohlsson, S. & Ohlsson, R. (1991). Expression of the human PDGF-B gene is regulated by both positively and negatively acting cell type-specific regulatory elements located in the first intron. *The EMBO Journal* **10**, 1365-1373.

Fuks, F., Burgers, W.A., Brehm, A., Hughes-Davies, L. & Kouzarides, T. (2000). DNA methyltransferase Dnmt1 associates with histone deacetylase activity. *Nat Genet* **24**, 88-91.

Fukushige, S. & Sauer, B. (1992). Genomic targeting with a positive-selection *lox* integration vector allows highly reproducible gene expression in mammalian cells. *Proc Natl Acad Sci USA* **89**, 7905-7909.

Gallie, D.R. (1998). Controlling gene expression in transgenics. *Current Opinion in Plant Biology* **1**, 166-172.

Garrick, D., Sutherland, H., Robertson, G. & Whitelaw, E. (1996). Variegated expression of a globin transgene correlates with chromatin accessibility but not methylation status. *Nucleic Acids Res.* **24**, 4902-4909.

Garrick, D., Fiering, S., Martin, D.I.K. & Whitelaw, E. (1998). Repeat-induced gene silencing in mammals. *Nat Genet* **18**, 56-59.

Gasser, S.M., Paro, R., Stewart, F. & Aasland, R. (1998). Epigenetic control of transcription. *Cellular and Molecular Life Sciences* **54**, 1-5.

Gordon, J.W., Scangos, G.A., Plotkin, D.J., Barbosa, J.A. & Ruddle, F.H. (1980). Genetic transformation of mouse embryos by microinjection of purified DNA. *Proc Natl Acad Sci USA* **77**, 7380-7384.

Gressel, J. (2000). Molecular biology of weed control. *Transgenic Res.* **9**, 355-382.

Gu, H., Marth, J.D., Orban, P.C., Mossmann, H. & Rajewsky, K. (1994). Deletion of a DNA polymerase β gene segment in T cells using cell type-specific gene targeting. *Science* **265**, 103-106.

Guo, F., Gopaul, D.N. & Van Duyne, G.D. (1997). Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse. *Nature* **389**, 40-46.

Gutiérrez-Adán, A., Maga, E.A., Behboodi, E., Conrad-Brink, J.S., Mackinlay, A.G., Anderson, G.B. & Murray, J.D. (1999). Expression of bovine b-lactoglobulin in the milk of transgenic mice. *Journal of Dairy Research* **66**, 289-294.

Guy, L.-G., Kothary, R. & Wall, L. (1997). Position effects in mice carrying a lac-Z transgene in *cis* with the β-globin LCR can be explained by a graded model. *Nucleic Acids Res.* **25**, 4400-4407.

Hamada, T., Sasaki, H., Seki, R. & Sakaki, Y. (1993). Mechanism of chromosomal integration of transgenes in microinjected mouse eggs - sequence analysis of genome-transgene and transgene-transgene junctions at two loci. *Gene* **128**, 197-202.

Hamilton, A.J. & Baulcombe, D.C. (1999). A Species of Small Antisense RNA in Posttranscriptional Gene Silencing in Plants. *Science* **286**, 950-952.

Hansen, R.S., Wijmenga, C., Stanek, A.M., Canfield, T.K., Weemaes, C.M.R. & Gartler, S.M. (1999). The *DNMT3B* DNA methyltransferase gene is mutated

in the ICF immunodeficiency syndrome. *Proc Natl Acad Sci USA* **96**, 14412-14417.

Hark, A.T., Schoenherr, C.J., Katz, D.J., Ingram, R.S., Levorse, J.M. & Tilghman, S.M. (2000). CTCF mediates methylation-sensitive enhancer-blocking activity at the *H19/Igf2* locus. *Nature* **405**, 486-489.

Harlow, E. & Lane, D. (1988). *Antibodies A laboratory manual.* Cold Spring Harbor Laboratory Publications New York, pp. 1-711.

Harris, S., Ali, S., Anderson, S., Archibald, A.L. & Clark, A.J. (1988). Complete nucleotide sequence of the genomic ovine β-lactoglobulin gene. *Nucleic Acids Res.* **16 (21)**, 10379-10380.

Harris, S., McClenaghan, M., Simons, J.P., Ali, S. & Clark, A.J. (1991). Developmental regulation of the sheep β-lactglobulin gene in transgenic mice. *Dev Genet* **12**, 299-307.

Heard, E., Mongelard, F., Arnaud, D. & Avner, P. (1999). Xist yeast artifical chromosome transgenes function as X-inactivation centers only in multicopy arrays and not as single copies. *Mol Cell Biol* **19**, 3156-3166.

Henchoz, S., De Rubertis, F., Pauli, D. & Spierer, P. (1996). The dose of a putative ubiquitin-specific protease affects position effect variegation in *Drosophila melanogaster*. *Mol Cell Biol* **16**, 5717-5725.

Henikoff, S. (1990). Position-effect variegation after 60 years. *Trends Genet.* **12**, 422-426.

Henikoff, S. (1996). Dosage-dependent modification of position-effect variegation in *Drosophila*. *BioEssays* **18**, 401-409.

Hennighausen, L. & Robinson, G.W. (1998). Think globally, act locally: the making of a mouse mammary gland. *Genes Dev.* **12**, 449-455.

Hoess, R.H. & Abremski, K. (1984). Interaction of the bacteriophage P1 recombinase Cre with the recombining site *loxP*. *Proc Natl Acad Sci USA* **81**, 1026-1029.

Hoess, R.H., Wierzbicki, A. & Abremski, K. (1986). The role of the *loxP* spacer region in P1 site-specific recombination. *Nucleic Acids Res.* **14**, 2287-2300.

Houdebine, L.M. (1997). Overview. In *Transgenic Animals: Generation and Use* (Houdebine, L.M., ed.), Harwood Academic Publishers, Amsterdam, pp. 1-6.

Houdebine, L.M. (2000). Transgenic animal bioreactors. *Transgenic Res.* **9**, 305-320.

Ibáñez, E., Folch, J.M., Vidal, F., Coll, A., Santaló, J., Egozcue, J. & Sánchez, A. (1997). Expression of caprine b-lactoglobulin in the milk of transgenic mice. *Transgenic Res.* **6**, 69-74.

Irelan, J.T. & Selker, E.U. (1996). Gene silencing in filamentous fungi: RIP, MIP and quelling. *Journal of Genetics* **75 (3)**, 313-324.

Jaenisch, R. & Mintz, B. (1974). Simian Virus 40 DNA sequences in DNA of helathy adult mice derived form preimplantation blastocysts injected with viral DNA. *Proc Natl Acad Sci USA* **71**, 1250-1254.

Jaenisch, R. (1976). Germ line integration and Mendelian transmission of the exogenous Molony leukemia virus. *Proc Natl Acad Sci USA* **73**, 1260-1264.

Jakowitsch, J., Papp, I., Moscone, E.A., van der Winden, J., Matzke, M.A. & Matzke, A.J.M. (1999). Molecular and cytogenetic characterization of a transgene locus that induces silencing and methylation of homologous promoter in *trans*. *The Plant Journal* **17**, 131-140.

James, R.M., Neil, C., Webster, J., Roos, S., Clark, A.J. & Whitelaw, C.B.A. (2000). Multiple copies of b-lactoglobulin promoter do not function as LCR. *Biochem.Biophys.Res.Comm.* **272**, 284-289.

James, T.C., Eissenberg, J.C., Craig, C., Dietrich, V., Hobson, A. & Elgin, S.C.R. (1989). Distribution patterns of HP1, a heterochromatin-associated nonhistone chromosomal protein of Drosophila. *European Journal of Cell Biology* **50**, 170-180.

Jasin, M., Moynahan, M.E. & Richardson, C. (1996). Targeted transgenesis. *Proc Natl Acad Sci USA* **93**, 8804-8808.

Jeddeloh, J.A., Bender, J. & Richards, E.J. (1998). The DNA methylation locus *DDM1* is required for maintance of gene silencing in *Arabidopsis*. *Genes Dev.* **12**, 1714-1725.

Jeddeloh, J.A., Stokes, T.L. & Richards, E.J. (1999). Maintance of genomic methylation requires a SWI2/SNF2-like protein. *Nat Genet* **22**, 94-97.

Jensen, S., Gassama, M.-P. & Heidmann, T. (1999). Taming of transposable elements by homology-dependent gene silencing. *Nat Genet* **21**, 209-212.

Jones, D.O., Cowell, I.G. & Singh, P.B. (2000). Mammalian chromodomain proteins: their role in genome organisation and expression. *BioEssays* **22**, 124-137.

Kühn, R. & Schwenk, F. (1997). Advances in gene targeting methods. *Current Opinion in Immunology* **9**, 183-188.

Keohane, A.M., Lavender, J.S., O'Neill, L.P. & Turner, B.M. (1998). Histone acetylation and X inactivation. *Dev Genet* **22**, 65-73.

Kilby, N.J., Snaith, M.R. & Murray, J.A.H. (1993). Site-specific recombinases: tools for genome engineering. *Trends Genet.* **9**, 413-421.

Kooter, J.M., Matzke, M.A. & Meyer, P. (1999). Listening to the silent genes: transgene silencing, gene regulation and pathogen control. *TIPS* **4**, 340-347.

Kornberg, R.D. & Lorch, Y.L. (1999). Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell* **98**, 285-294.

Lakso, M., Sauer, B., Mosinger, J.B., Lee, E.J., Manning, R.W., Yu, S.-H., Mulder, K.L. & Westphal, H. (1992). Targeted oncogene activation by site-specific recombination in transgenic mice. *Proc Natl Acad Sci USA* **89**, 6232-6236.

Lakso, M., Pichel, J.G., Gorman, J.R., Sauer, B., Okamoto, Y., Lee, E., Alt, F.W. & Westphal, H. (1996). Efficient *in vivo* manipulation of mouse genomic sequences at the zygote stage. *Proc Natl Acad Sci USA* **93**, 5860-5865.

Lallemand, Y., Luria, V., Haffner-Krausz, R. & Lonai, P. (1998). Maternally expressed PGK-Cre transgene as a tool for early and uniform activation of the Cre site-specific recombinase. *Transgenic Res.* **7**, 105-112.

Lamb, B.T. & Gearhart, J.D. (1995). YAC transgenics and the study of genetics and human disease. *Current Opinion in Genetics and Development* **5**, 342-348.

Lee, J.T., Davidow, L.S. & Warshawsky, D. (1999). *Tsix*, a gene antisense to *Xist* at the X-inactivation centre. *Nat Genet* **21**, 400-404.

Lee, J.T. & Lu, N. (1999). Targeted Mutagenesis of *Tsix* Leads to Nonrandom X Inactivation. *Cell* **99**, 47-57.

Li, E., Bestor, T.H. & Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**, 915-926.

Li, Z., Stark, G., Gotz, J., Rulicke, T., Muller, U. & Weissmann, C. (1996). Generation of mice with a 200-kb amyloid precursor protein gene deletion by Cre recombinase-mediated site-specific recombination in embryonic stem cells. *Proc Natl Acad Sci USA* **93**, 6158-6162.

Logan, J.S. (2000). Prospects for xenotransplantation. *Current Opinion in Immunology* **12**, 563-568.

Ludwig, D.L., Stringer, J.R., Wight, D.C., Doetschman, T.C. & Duffy, J.J. (1996). FLP-mediated site-specific recombination in microinjected murine zygotes. *Transgenic Res.* **5**, 385-395.

Luff, B., Pawlowski, L. & Bender, J. (1999). An Inverted Repeat Triggers Cytosine Methylation of Identical Sequences in *Arabidopsis*. *Molecular Cell* **3** , 505-511.

Lyko, F., Ramsahoya, B.H., Kashevsky, H., Tudor, M., Mastrangelo, M.-A., Orr-Weaver, T.L. & Jaenisch, R. (1999). Mammalian (cytosine-5) methyltransferases cause genomic DNA methylation and lethality in *Drosophila*. *Nat Genet* **23**, 363-366.

Lyon, M.F. (1998). X-Chromosome inactivation: a repeat hypothesis. *Cytogenetics and Cell Genetics* **80**, 133-137.

Mack, A., Sauer, B., Abremski, K. & Hoess, R. (1992). Stoichiometry of the Cre recombinase bound to the *lox* recombining site. *Nucleic Acids Res.* **20**, 4451-4455.

Maga, E.A. & Murray, J.D. (1995). Mammary gland expression of transgenes and the potential for altering the properties of milk. *Bio/Tech.* **13**, 1452-1457.

Marahrens, Y. (1999). X-inactivation by chromosomal pairing events. *Genes Dev.* **13**, 2624-2632.

Matzke, M.A. & Matzke, A.J.M. (1995). Homology-dependent gene silencing in transgenic plants: what does it really tell us? *Trends Genet.* **11**, 1-3.

Matzke, M.A., Matzke, A.J.M. & Eggleston, W.B. (1996). Paramutation and transgene silencing: a common response to invasive DNA? *TIPS* **1**, 382-388.

Matzke, M.A. & Matzke, A.J.M. (1998). Epigenetic silencing of plant transgenes as a consequence of diverse cellular defence reponses. *Cellular and Molecular Life Sciences* **54**, 94-103.

McBurney, M.W. (1999). Gene silencing in the development of cancer. *Experimental Cell Research* **248**, 25-29.

McCreath, K.J., Howcroft, J., Campbell, K.H.S., Colman, A., Schnieke, A.E. & Kind, A.J. (2000). Production of gene-targeted sheep by nuclear transfer from cultured somatic cells. *Nature* **405**, 1066-1069.

McFarlane, M. & Wilson, J.B. (1996). A model for the mechanism of precise integration of a microinjected transgene. *Transgenic Res.* **5**, 171-177.

Mehtali, M., LeMeur, M. & Lathe, R. (1990). The methylation-free status of a housekeeping transgene is lost at high copy number. *Gene* 179-184.

Mercier, J.C., Gaye, P., Soulier, S., Huedelahaie, D. & Vilotte, J.L. (1985). Construction and identification of recombinant plasmids carrying cDNAs coding for ovine alpha-S1-casein, alpha-S2-casein, beta-casein, kappa-casein and beta-lactoglobulin - nucleotide sequence of alpha-S1-casein cDNA. *Biochimie* **67**, 959-971.

Metzger, D. & Feil, R. (1999). Engineering the mouse genome by site-specific recombination. *Current Opinion in Biotechnology* **10**, 470-476.

Mintz, B. & Bradl, M. (1991). Mosaic expression of a tyrosinase fusion gene in albino mice yields a heritable striped coat color pattern in transgenic homozygotes. *Proc Natl Acad Sci USA* **88**, 9643-9647.

Mochizuki, A., Takeda, Y. & Iwasa, Y. (1996). The evoution of Genomic Imprinting. *Genetics* **144**, 1283-1295.

Moore, T. & Haig, D. (1991). Genomic imprinting in mammalian development -a parental tug-of-war. *Trends Genet.* **7**, 45-49.

Morgan, H.D., Sutherland, H.G.E., Martin, D.I.K. & Whitelaw, E. (1999). Epigenetic inheritance at the agouti locus in the mouse. *Nat Genet* **23**, 314-318.

Muller, H.J. (1930). Types of visible variations induced by X-rays in Drosophila. *Journal of Genetics* **12**, 299-334.

Nadeau, J.H. (2001). Modifier genes in mice and humans. *Nature Reviews Genetics* **2**, 165-174.

Nakajima, T., Uchida, C., ANderson, S.F., Lee, C.G., Hurwitz, J., Parvin, J.D. & Montminy, M. (1997). RNA helicase a mediates association of CBP with RNA polymerase II. *Cell* **90**, 1107-1112.

Namciu, S.J., Blochlinger, K.B. & Fournier, R.E.K. (1998). Human Matrix Attachment Regions Insulate Transgene Expression from Chromsomal Position Effects in Drosophila melanogaster. *Mol Cell Biol* **18**, 2382-2391.

Nan, X., Ng, H.-H., Johnson, C.A., Laherty, C.D., Turner, B.M., Eisenman, R.N. & Bird, A. (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* **393**, 386-389.

Nicoll, C.S., & Tucker,A. (1965). Estimates of parenchymal, stromal, and lymph node deoxyribonucleic acid in mammary glands of C3H/Crgl/2 mice. *Life Sciences* **4**, 993-1001.

Old, R.W. & Primrose, S.B. (1994). The impact of recombinant DNA technology: the generation of novelty. In *Principles of Gene Manipulation* (Old, R.W. and Primrose, S.B., eds.), Blackwell Science Ltd, Oxford, pp. 369-407.

Orban, P.C., Chui, D. & Marth, J.D. (1992). Tissue- and site-specific DNA recombination in transgenic mice. *Proc Natl Acad Sci USA* **89**, 6861-6865.

Pal-Bhadra, M., Bhadra, U. & Birchler, J.A. (1997). Cosuppression in Drosophila: Gene silencing of *Alcohol dehydrogenase* by *white-Adh* transgenes is *Polycomb* dependent. *Cell* **90**, 479-490.

Pal-Bhadra, M., Bhadra, U. & Birchler, J.A. (1999). Cosuppression of nonhomologous transgenes in Drosophila involves mutually related endogenous sequences. *Cell* **99**, 35-46.

Palmiter, R.D., Chen, H.Y. & Brinster, R.L. (1982). Differential regulation of metallothionein-thymidine kinase fusion genes in transgenic mice and their offspring. *Cell* **29**, 701-710.

Palmiter, R.D. & Brinster, R.L. (1985). Transgenic Mice. *Cell* **41**, 343-345.

Paro, R. & Hogness, D.S. (1991). The polycomb protein shares a homologous domain with a heterochromatin-associated protein of *Drosophila*. *Proc Natl Acad Sci USA* **88**, 263-267.

Pervaiz, S. & Brew, K. (1985). Homology of Beta-lactoglobulin, serum retinol-binding protein, and protein HC. *Science* **228**, 335-337.

Peterson, C.L. & Workman, J.L. (2000). Promoter targeting and chromatin remodeling by the SWI/SNF complex. *Current Opinion in Genetics and Development* **10**, 187-192.

Petters, R.M. & Sommer, J.R. (2000). Transgenic animals as models for human disease. *Transgenic Res.* **9**, 347-351.

Powers, J.A. & Eissenberg, J.C. (1993). Overlapping domains of the heterochromatin-associated protein HP1 mediate nuclear localisation and heterochromatin binding. *The Journal of Cell Biology* **120**, 291-299.

Pursel, V.G. & Rexroad Jr., C.E. (1993). Status of research with transgenic farm animals. *J.Animal Sci.* **71 (Suppl. 3)**, 10-19.

Pursel, V.G. & Rexroad, C.E. (1993). Status of research with transgenic farm animals. *J.Animal Sci.* **71**, 10-19.

Rakyan, V.K., Preis, J., Morgan, H.D. & Whitelaw, E. (2001). The marks, mechanisms and memory of epigenetic states in mammals. *Biochem J.* **356**, 1-10.

Ramirez-Solis, R., Liu, P., Bradley,A. (1995). Chromosome engineering in mice. *Nature* **378 (6558),** 720-724.

Ramírez, A., Milot, E., Ponsa, I., Marcos-Gutiérrez, C., Page, A., Santos, M., Jorcano, J. & Vidal, M. (2001). Sequence and chromosomal context effects on variegated expression of Keratin 5/lacZ constructs in stratified epithelia of transgenic mice. *Genetics* **158**, 341-350.

Ratcliff, F., Harrison, B.D. & Baulcombe, D.C. (1997). A similarity between viral defense and gene silencing in plants. *Science* **276**, 1558-1560.

Reik, W. & Walter, J. (2001). Genomic imprinting: Parental influence on the genome. *Nature Reviews Genetics* **2**, 21-32.

Robertson, G., Garrick, D., Wu, W., Kearns, M., Martin, D.I.K. & Whitelaw, E. (1995). Position-dependent variegation of globin transgene expression in mice. *Proc Natl Acad Sci USA* **92**, 5371-5375.

Rudolph, N.S. (1999). Biopharmaceutical production in transgenic livestock. *Trends Biotech.* **17**, 367-374.

Sabl, J.F. & Henikoff, S. (1996). Copy number and orientation determine the susceptibility of a gene to silencing by nearby heterochromatin in Drosophila. *Genetics* **142**, 447-458.

Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989). *Molecular cloning: a laboratory manual.* Cold Spring Harbor Library Press Cold Spring Harbor, New York,

Sapino, A., Macri, L., Gugliotta, P., & Bussolati, G. (1990). Immunocytochemical identification of proliferating cell types in mouse mammary gland. *J Histochem Cytochem* **38 (11),** 1541-1547.

Sapino, A., Macri, L., Tonda, L., & Bussolati, G. (1993). Oxytocin enhances myoepithelial cell differentiation and proliferation in the mouse mammary gland. *Endocrinology* **133 (2),** 838-842.

Sauer, B. (1987). Functional expression of the *cre-lox* site-specific recombination system in the yeast *Saccharomyces cerevisiae. Mol Cell Biol* **7,** 2087-2096.

Sauer, B. & Henderson, N. (1989). Cre-stimulated recombination at *loxP*-containing DNA sequences placed into the mammalian genome. *Nucleic Acids Res.* **17,** 147-161.

Sauer, B. (1993). Manipulation of transgenes by site-specific recombination: Use of Cre recombinase. *Methods in Enzymology* **225,** 890-900.

Schedl, A., Montoliu, L., Kelsey, G. & Schütz, G. (1993). A yeast artificial chromosome covering the tyrosinase gene confers copy number-dependent expression in transgenic mice. *Nature* **362,** 258-261.

Schiebel, W., Pelissier, T., Riedel, L., Thalmeir, S., Schiebel, R., Kempe, D., Lottspeich, F., Sanger, H.L. & Wassenegger, M. (1998). Isolation of an RNA-directed RNA polymerase-specific cDNA clone from tomato. *Plant Cell* **10,** 2087-2101.

Schmidt, E.E., Taylor, D.S., Prigge, J.R., Barnett, S. & Capecchi, M.R. (2000). Illegitimate Cre-dependent chromosome rearrangements in transgenic mouse spermatids. *Proc Natl Acad Sci USA* **97,** 13702-13707.

Scrable, H. & Stambrook, P.J. (1999). A genetic program for deletion of foreign DNA from the mammalian genome. *Mutation Research* **429,** 225-237.

Selbert, S., Bentley, D.J., Melton, D.W., Rannie, D., Lourenco, P., Watson, C.J. & Clarke, A.R. (1998). Efficient BLG-Cre mediated gene deletion in the mammary gland. *Transgenic Res.* **7**, 387-396.

Shani, M., Barash, I., Nathan, M., Ricca, G., Searfoss, G.H. & Dekel, I. (1992). Expression of human serum albumin in the milk of transgenic mice. *Transgenic Res.* **1**, 195-208.

Sharpe, J.A., Chan-Thomas, P.S., Lida, J., Ayyub, H., Wood, W.G. & Higgs, D.R. (1992). Analysis of the human α globin upstream regulatory element (HS -40) in transgenic mice. *The EMBO Journal* **11**, 4565-4572.

Siegfried, Z. & Cedar, H. (1997). DNA methylation: A molecular lock. *Curr.Biol.* **7**, R305-R307

Simons, J.P., McClenaghan, M. & Clark, A.J. (1987). Alteration of the quality of milk by expression of sheep β-lactoglobulin in transgenic mice. *Nature* **328**, 530-532.

Southern, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology* **98**, 503-517.

Stief, A., Winter, D.M., Strätling, W.H. & Sippel, A.E. (1989). A nuclear DNA attachment element mediates elevated and position-dependent gene activity. *Nature* **341**, 343-345.

Struhl, K. (1998). Histone acetylation and transcriptional regulatory mechanisms. *Genes and Development* **12**, 599-606.

Sutherland, H.G.E., Kearns, M., Morgan, H.D., Headley, A.P., Morris, C., Martin, D.I.K. & Whitelaw, E. (2000). Reactivation of heritably silenced gene expression in mice. *Mammalian Genome* **11**, 347-355.

Thorvaldsen, J.L., Duran, K.L. & Bartolomei, M.S. (1998). Deletion of the *H19* differentially methylated domain results in loss of imprinted expression of *H19* and *Igf2*. *Genes and Development* **12**, 3693-3702.

Thyagarajan, B., Guimarães, M.J., Groth, A.C. & Calos, M.P. (2000). Mammalian genomes contain active recombinase recognition sites. *Gene* **244**, 47-54.

Townes, T.M., Lingrel, J.B., Chen, H.Y., Brinster, R.L. & Palmiter, R.D. (1985). Erythroid-specific expression of human β-globin genes in transgenic mice. *The EMBO Journal* **4**, 1715-1723.

Tucker, H.A. (2000). Hormones, Mammary Growth and Lactation:a 41-Year Perspective. *J.Dairy Sci.* **83**, 874-884.

Turner, B.M. (1998). Histone acetylation as an epigenetic determinant of long-term transcriptional competence. *Cellular and Molecular Life Sciences* **54**, 21-31.

Tycko, B. (2000). Epigenetic gene silencing in cancer. *The Journal of Clinical Investigation* **105 (4)**, 401-407.

Urieli-Shoval, S., Gruenbaum, Y., Sedat, J. & Razin, A. (1982). The absence of detectable methylated bases in *Drosophila melanogaster* DNA. *FEBS Lett.* **146**, 148-152.

Voinnet, O., Vain, P., Angell, S. & Baulcombe, D.C. (1998). Systematic Spread of Sequence-Specific Transgene RNA Degradation in Plants is Initiated by Localized Introduction of Ectopic Promoterless DNA. *Cell* **95**, 177-187.

Wagner, K.U., Wall, R.J., St-Onge, L., Gruss, P., Wynshaw-Boris, A., Garrett, L., Li, M., Furth, P.A. & Hennighausen, L. (1997). Cre-mediated gene deletion in the mammary gland. *Nucleic Acids Res.* **25 (21)**, 4323-4330.

Wall, R.J. & Seidel Jr, G.E. (1992). Transgenic farm animals - A critical analysis. *Theriogenology* **38**, 337-357.

Wall, R.J. & Burdon, T.G. (1997). The fate of microinjected genes in preimplantation embryos. In *Transgenic Animals: Generation and Use* (Houdebine, L.M., ed.), Harwood Academic Publishers, Amsterdam, pp. 215-217.

Wallace, H., Ansell, R., Clark, J. & McWhir, J. (2000). Pre-selection of integration sites imparts repeatable transgene expression. *Nucleic Acids Res.* **28**, 1455-1464.

Webster, J., Wallace, R.M., Clark, A.J. & Whitelaw, C.B.A. (1995). Tissue-specific, temporally regulated expression mediated by the proximal ovine β-lactoglobulin promoter in transgenic mice. *Cell.Mol.Biol.Res.* **41**, 11-16.

Whitelaw, C.B.A., Harris, S., McClenaghan, M., Simons, J.P. & Clark, A.J. (1992). Position-independent expression of the ovine β-lactoglobulin gene in transgenic mice. *Biochem J.* **286**, 31-39.

Whitelaw, C.B.A., Springbett, A.J., Webster, J. & Clark, A.J. (1993). The majority of G(0) transgenic mice are derived from mosaic embryos. *Transgenic Res.* **2**, 29-32.

Whitelaw, C.B.A. (1995). Regulation of ovine β-lactoglobulin gene expression during the first stage of lactogenesis. *Biochem.Biophys.Res.Comm.* **209**, 1089-1093.

Whitelaw, C.B.A. (1996). Hormonal influences on β-lactoglobulin transgene expression inferred from chromatin structure. *Biochem.Biophys.Res.Comm.* **224**, 121-125.

Whitelaw, C.B.A. (2000). Nucleosome organisation of the beta-lactoglobulin gene - Transcription complex formation. *Biology of the Mammary Gland: Advances in Experimental Medicine and Biology* **480**, 147-153.

Whitelaw, C.B.A., Grolli, S., Accornero, P., Donofrio, G., Farini, E. & Webster, J. (2000). Matrix attachment region regulates basal β-lactoglobulin transgene expression. *Gene* **244**, 73-80.

Wilmut, I., Schnieke, A.E., McWhir, J., Kind, A.J. & Campbell, K.H.S. (1997). Viable offspring derived from fetal and adult mammalian cells. *Nature* **385**, 810-813.

Wolffe, A.P. (1997). Transcription control: Repressed repeats express themselves. *Curr.Biol.* **7**, R796-R798

Wolffe, A.P., Jones, P.L. & Wade, P.A. (1999). DNA methylation. *Proc Natl Acad Sci USA* **96**, 5894-5896.

Wu, C.-T. & Morris, J.R. (1999). Transvection and other homology effects. *Current Opinion in Genetics and Development* **9**, 237-246.

Yoder, J.A., Walsh, C.P. & Bestor, T.H. (1997). Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet.* **13**, 335-340.

Yudkovsk, N., Logie, C., Hahn, S. & Peterson, C.L. (1999). Recruitment of the SWI/SNF chromatin remodelling complex by transcriptional activators. *Genes and Development* **13**, 2369-2374.

Yull, F.E., Binas, B., Harold, G., Wallace, R. & Clark, A.J. (1997). Transgene rescue in the mammary gland is associated with transcription but does not require translation of BLG transgenes. *Transgenic Res.* **6**, 11-17.